ABSTRACT

Many bacteria able to degrade *n*-alkanes have been isolated, but only a few *n*-alkane oxidizing systems have been characterized so far. Furthermore, genes involved in degradation of *n*-alkanes longer than C_{16} have not been reported prior to 2007. The Gram positive GC-rich *Gordonia* sp. strain *SoCg*, isolated in Sicily from an accidentally contaminated beach, is able to grow on long chain *n*-alkanes up to C_{36}. The aim of my project was to investigate its alkane oxidation system(s). The kinetic of growth on two model *n*-alkanes (HC) hexadecane (liquid) and triacontane (solid) showed that rapid growth phase corresponds to a rapid HC degradation of both alkanes. The research of the catabolic genes revealed that this strain carries a single copy of a non-heme membrane bound alkane hydroxylase encoding gene (*alkB*), and sequence analysis of its flanking region revealed other five consecutive open reading frames (ORFs) which were designated Orf1, RubA3, RubA4, RubB and AlkU, according to the sequence homology with other known *alk* clusters. Expression of this genes was induced by both liquid and solid *n*-alkanes. The *alkB* gene was successfully expressed in *Escherichia coli* BL21, obtaining strain BL21-AH and *Streptomyces coelicolor* M145, obtaining strain M145-AH and 1-hexadecanol was detected by GC-MS analysis in both hosts as oxidative metabolic intermediate of *n*-hexadecane. Finally, the inability of a *Gordonia* SoCg *alkB* disruption mutant to grow on solid HC demonstrated, for the first time, the role of the gordonial *alkB* in catabolism of long chain *n*-alkanes.

Production of storage compounds during growth on HC was also explored.. A relationship between different chain length *n*-alkanes as substrate and neutral lipids (mono/di/tri-acyl-glycerols, wax esters) was detected using thin layer chromatography (TLC) analysis. Moreover, SPME GC-MS analysis showed the production of high value hexadecyl-hexadecanoic acid (WE). A putative acyltransferase encoding gene, involved in neutral lipids biosynthesis, was PCR-detected using degenerated primers.

Since, poor knowledge about physiological response to *n*-alkanes metabolism in Gram positive bacteria exist to date and growth ability on HC was acquired by strain M145-AH a proteomic approach was carried out to understand bacterial adaptation to HC substrates. Interestingly, differentially 2D-DIGE analysis showed that M145-AH modulates the expression of proteins involved in central carbon metabolism, fatty acid oxidation and biosynthesis. Surprisingly, *n*-hexadecane proteome profile was not much different from the proteome obtained on glucose, indicating an extreme metabolic versatility for *Streptomyces*. This global analysis could provide proteomic platform to investigate mechanisms underly bacterial HC bioconversion.

Introduction

metabolism of n-alkanes in Gordonia SoCg

Introduction

Alkanes are saturated, linear hydrocarbons whose chain length can vary from one (in methane) to more than fifty carbon atoms. Alkanes constitute about 20 to 50% of crude oil, depending on the source of the oil. However, also living organisms, such as bacteria, plants, and some animals, produce them. As a result, alkanes are widespread in nature, and many microorganisms have evolved enzymes to use them as a carbon source. Alkanes, however, are chemically inert and must be activated before they can be metabolized. In the presence of oxygen, activation is usually achieved by oxidation of one of the terminal methyl groups to generate the corresponding primary alcohol by alkane hydroxylases (Rojo 2005).

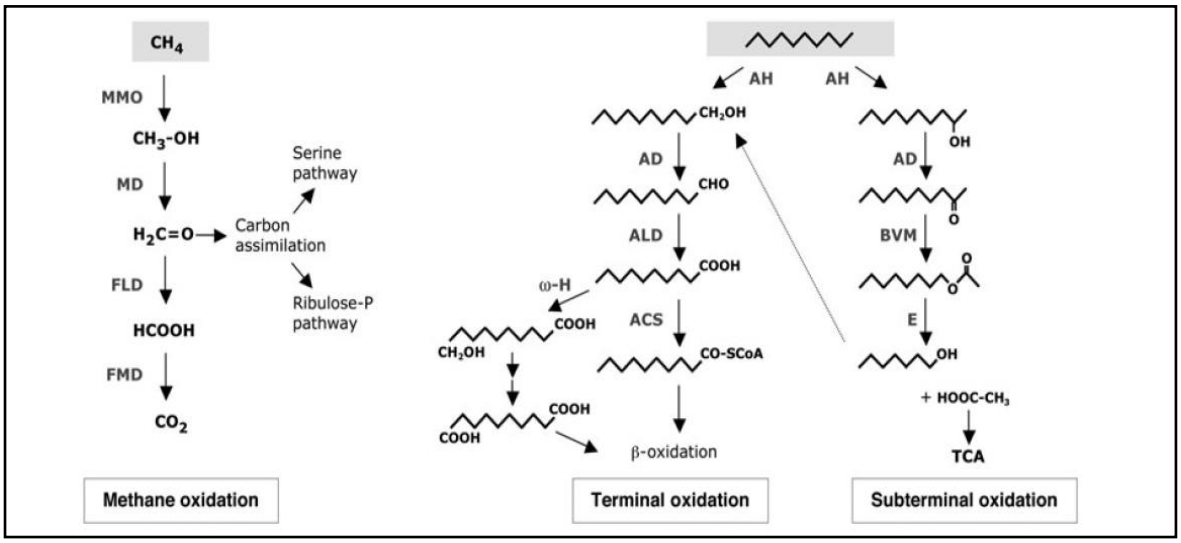

Fig.1 Aerobic pathways for the degradation of methane (left), and of larger *n*-alkanes by terminal and subterminal oxidation (right). Initial activation of the alkane molecule requires O2 as a reactant. MMO, methane monooxygenase; MD, methanol dehydrogenase, FLD, formaldehyde dehydrogenase; FMD, formate dehydrogenase. AH, alkane hydroxylase; AD, alcohol dehydrogenase; ALD, aldehyde dehydrogenase; ACS, acyl-CoA synthetase; w-H, w-hydroxylase; BVM, Baeyer–Villiger monooxygenase; E, esterase; TCA, tricarboxylic acids cycle.

Bacterial oxidation of *n*-alkanes is a very common phenomenon in soil and water and is a major process in geochemical terms: the estimated amount of alkanes that is recycled per year amounts to several million tons from natural oil seepage and oil spills alone. Although many microorganisms are capable of degrading aliphatic hydrocarbons and they are readily isolated from contaminated and non contaminated sites, relatively little is known about the genetic characteristics of their alkane-degradative systems.

Through intensive investigations on hydrocarbons petroleum microbiological degradation, a wide variety of microbial processes have been re-examined. Examples include, unique enzymes related to hydrocarbon degradation (Li et al., 2002), which can be used in the biotransformation of useful compounds; a variety of bio-emulsifiers (Toren et al., 2001; Toren et al., 2002) which are produced during microbial assimilation of hydrocarbons; and some cell-reserve materials of hydrocarbon-utilizers (Alvarez et al., 2000; Ishige et al., 2002). The recalcitrant liquid and solid *n*-alkane are oxidized to long chain fatty alcohol and finally, completely metabolized via oxidative process to fatty acid needed to support growth of the cell. It was demonstrated (Ishige et al., 2003) that part of the 1-alkanol is esterified with fatty acid CoA thioesters by the enzyme WS/DGAT (Wältermann et al., 2007). This reaction generate high energy value molecules, that are packed in internal vesicles as storage compounds. This phenomenon was descrive in *Acinetobacter* and in *Alcanivorax* strains and also in the oleaginous bacteria such as *Rhodococcus*. Bacteria differ in the type of reserve material that they accumulate, examples include glycogen, polyhydroxyalkanoates, triacylglycerols and polyphosphate (Finnerty et al., 2000). Uniquely, some strains of *Acinetobacter* produce and accumulate wax esters in enormous amounts from *n*-alkanes or long-chain alkanols under nitrogen-limited conditions.

The molecular and biochemical basis of microbial behavior and physiological responses to hydrocarbons and the impact of these responses on bioremediation have been neglected

until recently (van Hamme et al., 2003). Relatively speaking, the metabolic pathways driving the activation of hydrocarbons into central metabolic pathways are well understood, while behaviors and responses are not appreciated beyond a general observational level. However, these phenomena are essential for allowing hydrocarbon-metabolizing organisms to avoid toxic effects, to access poorly soluble substrates, and, in some cases, to bring very large substrates into the cell.

The genus *Gordonia*

The actinomycete genus *Gordonia* has attracted much interest in recent years for a variety of reasons. Most species were isolated due to their abilities to degrade xenobiotics, environmental pollutants, or otherwise slowly biodegradable natural polymers as well as to transform or synthesize possibly useful compounds. The variety of chemical compounds being transformed, biodegraded, and synthesized by gordoniae makes these bacteria potentially useful for environmental and industrial biotechnology. However, because some gordoniae are opportunistic pathogens, their application in the environment may be restricted in some cases.

Gordonia species have been isolated from various native biotopes such as soil or mangrove rhizosphere, from extensively industrially influenced habitats such as oil-producing wells or hydrocarboncontaminated soil, from artificial sources such as wastewater treatment bioreactors or biofilters, and from diseased humans.

The genus *Gordonia* belongs phylogenetically to the suborder *Corynebacterineae*, the mycolic acid group within the order Actinomycetales (Stackebrandt, et al., 1997), and its classification has changed drastically in recent years.

Incorporation of the long aliphatic chains of the mycolic acids into the cell wall is associated with hydrophobicity and surface adhesion (Bendinger et al., 1993) and may play a role in the

degradation of hydrophobic pollutants. Examples of this ability are the adhesive growth of several *Gordonia* strains during the biodegradation of rubber materials (Arenskötter, et al., 2001; Linos, et al., 2000) and the utilization of hydrophobic hydrocarbons by many species of this genus (Xue et al., 2003).

In 2004, the *n*-decane degrader *Gordonia* TF6 was isolated from the soil on a tennis court in Fujisawa, Kanagasawa, Japan (Fuji et al., 2004). An *alkB* cluster, like those found in *n*-alkane utilizing gram positive strain, was characterized for the first time in a strain of this genus. This cluster was heterologously expressed in *E.coli* and the corresponding 1-alchol of *n*-alkane bioconversion up to C_{13} was observed, indicating that this genomic region is the minimum component for *n*-alkane oxidation system.

Recently, *Gordonia sp.* SoCg strain was isolated from long term petroleum contaminated beach Sicily (Quatrini et al., 2008). It is able to growth on liquid and solid *n*-alkanes as sole carbon source, indicating an intersting metabolic versatility. An *alkB* gene like that of phylogenetically related *Gordonia* TF6 was found and partially sequenced. This thesis is actually the first report showing the ability of strain of this genus to degrade *n*-alkanes longer than C_{14}.

CHAPTER I

An alkane hydroxylase system of Gordonia sp. strain SoCg is involved in degradation of solid n-alkanes.

1. Alkane hydroxylases in Gram negative and Gram positive bacteria

Until recently, only the alkane-degradative genes of a small number of gram-negative bacteria, namely, *Psudomonas* and *Acinetobacter*, have been described in detail. Of these, the *alk* system found in *Psudomonas putida* GPo1, which degrades C_5 to C_{12} *n*-alkanes, remains the most extensively characterized alkane hydroxylase system (van Beilen et al., 2001). Much less is known about the alkane-degradative systems of Gram-positive bacteria. A putative alkane monooxygenase gene has been identified in the genome sequence of *Mycobacterium tubercolosis* H37Rv (Cole et al., 1998), while other *alkB* homologs were amplified from *Rhodococcus erythropolis* NRRL B-16531 (Smits et al., 1999) and in *Nocardiodes* sp. strain CF8 (Hamamura et al., 2001).

The *M. tubercolosis alkB* homologs could be functionally expressed in an *alkB* knockout derivative of *Pseudomonas fluorescens* CHA0 and in *P.putida* GPo1 and were shown to oxidize alkanes ranging from C_{10} to C_{16} (Smits et al., 2002).

Four alkane monoxygenase homologs (two as part of alkane gene clusters and two occurring as separate genes) were identified in two closely related *Rhodococcus* spp. and analysed by functional heterologous expression in *E. coli* and *Pseudomonas spp*. (Whyte et al., 2002). Similarly genes, encoding an AH system (alkane 1-monoxygenase, rubredoxins and rubredoxin reductase), from *Gordonia TF6* were cloned, sequenced and expressed in *E.coli*, where they resulted to be the minimum component to confer alkane hydroxylase activity in this strain on *n*-alkanes up to C_{12} (Fujii et al., 2004).

Alkanes longer than C_{16} support growth of many microorganisms, but the identity of enzymes involved in their oxidation is still an open question, especially in Gram positive actinobacteria (van Beilen and Funhoff 2007).

2. Aerobic bacterial degradation of *n*-alkanes

During the past decades, research related to alkane degradation has focused on the identification and characterization of enzymes involved in the initial step of aerobic bacterial catabolic pathways. In most described cases, the *n*-alkane is oxidized to the corresponding primary alcohol by substrate-specific terminal monooxygenases/hydroxylases. However, subterminal oxidation has also been described both for long-chain *n*-alkane substrates up to C_{16} (Whyte et al. 1998) and for *n*-alkanes of shorter chain lengths (Ashraf et al. 1994; Sullivan et al. 1998). The class of alkane hydroxylases involved in bacterial aerobic *n*-alkane metabolism has recently been reviewed in detail by van Beilen and Funhoff (2007). In that review, two unrelated classes of enzymes for long-chain *n*-alkane oxidation were proposed: (**1**) the class of bacterial particulate alkane hydroxylases (pAHs); (**2**) the class of cytochrome-P450-related enzymes in both yeasts and bacteria, e.g., bacterial CYP153 enzymes. The first class of integral membrane non-heme diiron monooxygenases of the AlkB-type allows a wide range of Proteobacteria and Actinomycetales to grow on *n*-alkanes with carbon chain lengths from C_5 to C_{16}.

After initial oxidation of the *n*-alkane, the corresponding alcohol is subsequently oxidized further by alcohol dehydrogenase and aldehyde dehydrogenase to the corresponding aldehyde and carboxylic acid, respectively. The carboxylic acid then serves as a substrate for acyl-CoA synthetase, and the resulting acyl-CoA enters the β- oxidation pathway (Fig. 2-3)

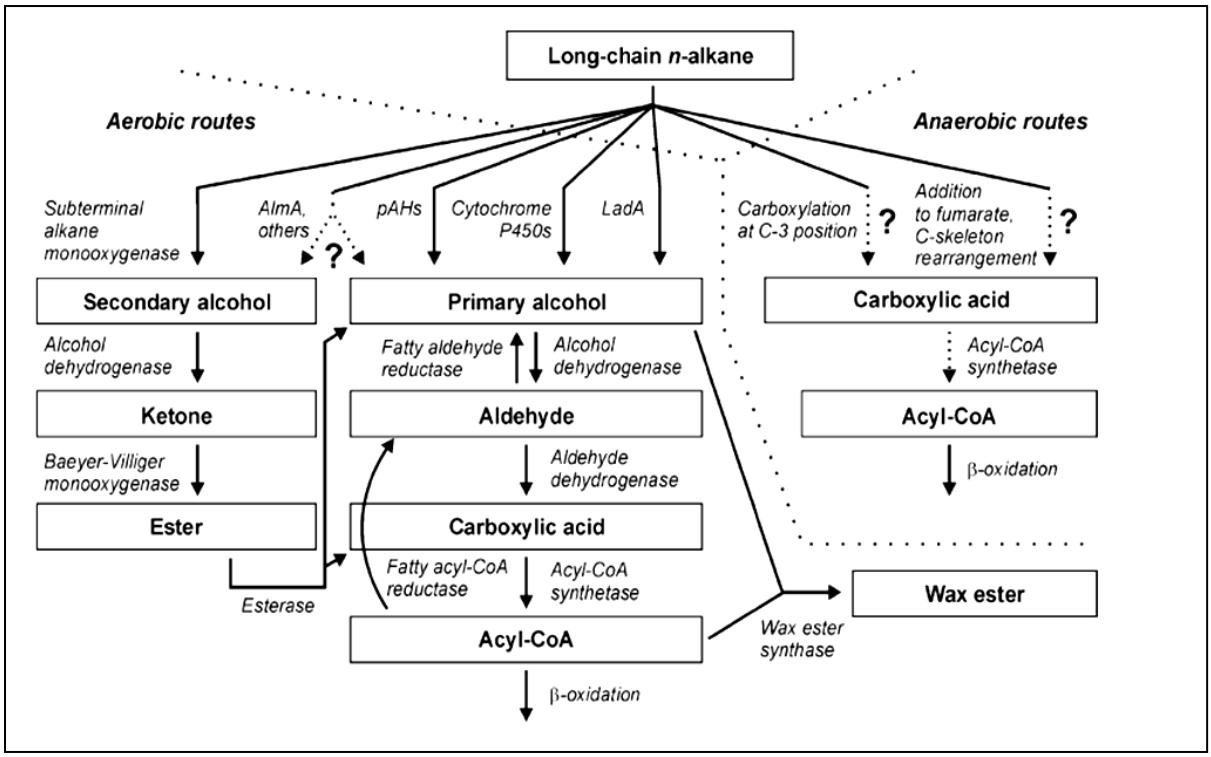

Figure 2 Pathways for aerobic and anaerobic bacterial degradation of long-chain *n*-alkanes and for synthesis of wax esters. Dotted arrows represent suggested metabolic routes.

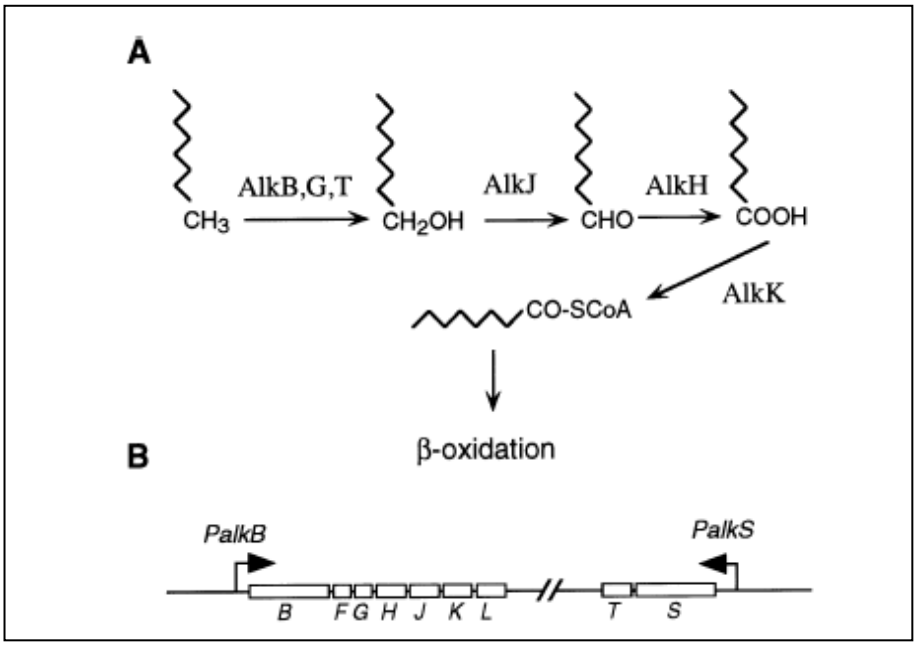

Figure 3. The alkane oxidation pathway in *Pseudomonas oleovorans GPo1*. Assimilation of medium chain length n-alkanes by the enzymes of the OCT plasmid involves a sequential oxidation of a terminal methyl group by an alkane hydroxylase (genes alkB,G,T), an alcohol dehydrogenase (alkJ ene) and an aldehyde dehydrogenase (alkH gene). After being activated by an acyl- oA synthetase (alkK gene), the fatty acids enter the b-oxidation cycle. B. The genes are grouped in two clusters; alkS codes for a transcriptional regulator, AlkS, which activates expression of the pathway in the presence of alkanes (adapted from van Beilen et al., 1994).

2.1. The AlkB family of alkane hydroxylases

AlkB-type enzymes function in complex with two electron transfer proteins, a dinuclear iron rubredoxin, and a mononuclear iron rubredoxin reductase channeling electrons from NADH to the active site of the alkane hydroxylase (van Beilen et al. 2003). Rubredoxin reductase, via its cofactor FAD, transfers electrons from NADH to the rubredoxin, which in turn transfers the electrons to AlkB. (Fig. 4). The active site includes four His-containing sequence motives that are conserved in other hydrocarbon monooxygenases, and which chelate two iron atoms (van Beilen *et al*., 1992; Shanklin *et al*., 1994). The diiron cluster allows the O_2-dependent activation of the alkane through a substrate radical intermediate (Shanklin *et al*., 1997; Austin *et al*., 2000; Bertrand *et al*., 2005). One of the O_2 atoms is transferred to the terminal methyl group of the alkane, rendering an alcohol, while the other one is reduced to H_2O by electrons transferred by the rubredoxin.

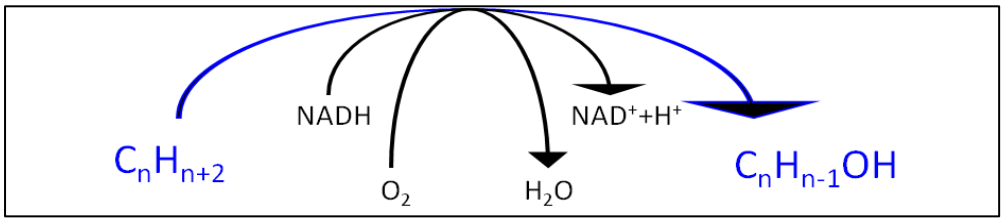

Figure 4. Alkanes are chemically inert and must be activated before they can be metabolized. In the presence of oxygen, activation is usually achieved by oxidation of one of the terminal methyl groups to generate the corresponding primary alcohol .

2.1.1. Molecular structure of AlkB

The *P. putida* GPo1 AlkB alkane hydroxylase can oxidize propane, *n*-butane (Johnson and Hyman, 2006), as well as C_5 to C_{13} alkanes (van Beilen *et al.*, 2005). All these alkanes can also support growth. Methane, ethane, or alkanes longer than C_{13}, are not oxidized. AlkB has been proposed to contain a deep hydrophobic pocket formed by the six transmembrane helices; the alkane molecule should slide into this pocket until the terminal methyl group is correctly positioned relative to the His residues that chelate the iron atoms (van Beilen *et al.*, 2005b) (Fig. 5). Amino acids with bulky side-chains protruding into the hydrophobic pocket can impose a limit to the size of the alkane molecule that can slide into the pocket and still allow a proper alignment of the terminal methyl group with the catalytic His residues. Substitution of these amino acids by residues with less bulky side-chains allows larger alkanes to fit in place into the hydrophobic pocket.

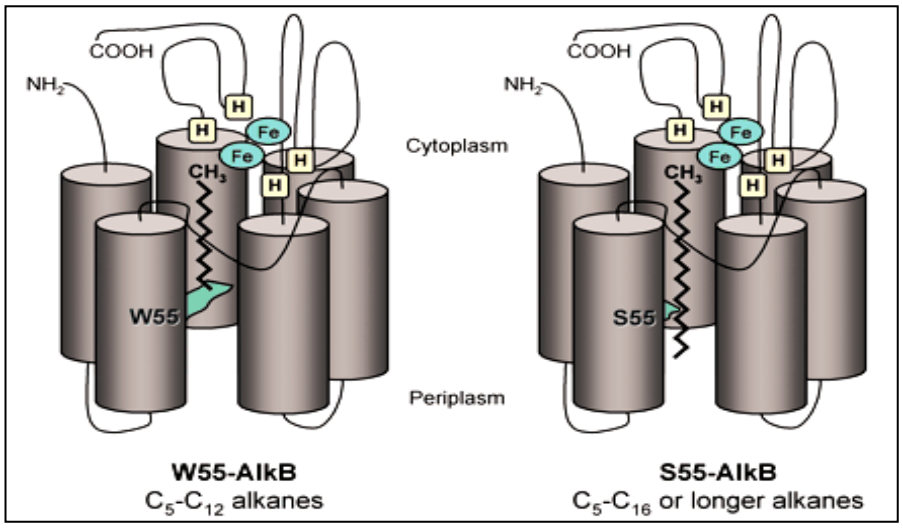

Figure 5. Model for the AlkB alkane hydroxylase. The enzyme is proposed to have six transmembrane helices arranged in a hexagonal distribution that would define a long, hydrophobic pocket into which the linear alkane molecule can slip. The four histidine clusters (H) believed to bind the two iron atoms (Fe) would lie on the cytoplasmic side. In *P. putida* GPo1 AlkB, residue W55 lies in transmembrane helix 2 and extends its bulky arm towards the hydrophobic pocket (left). This hampers the proper insertion of alkanes longer than C_{13}. The replacement of W55 by a serine residue (S55) (right), which has a shorter arm, allows longer alkanes to enter the pocket without impeding the proper alignment of the terminal methyl group relative to the histidine clusters. Sizes are not to scale. The soluble components of the alkane hydroxylase complex, rubredoxin and rubredoxin reductase, are not shown. The model is modified from reference 20 with permission.

2.2. Cytochrome P450 alkane hydroxylases

Cytochromes P450 are hemoproteins that hydroxylate a large number of compounds. They are ubiquitous among all kingdoms of life and can be grouped in more than 100 families on the basis of sequence similarity. Several bacterial strains that degrade C_5–C_{10} alkanes contain alkane hydroxylases that belong to a distinct family of soluble cytochrome P450 monooxygenases. The first member characterized was CYP153A1 from *Acinetobacter* sp. EB104 (Maier et al., 2001), but similar enzymes have been found in diverse strains of mycobacteria, rhodococci and proteobacteria (van Beilen et al., 2005; 2006; Sekine et al., 2006). These cytochromes P450 require a ferredoxin and a ferredoxin reductase that transfer electrons from NAD(P)H to the cytochrome. The cytochrome P450 from *Mycobacterium* sp. HXN-1500 was purified and shown to hydroxylate C_6–C_{11} alkanes to 1-alkanols with high affinity and regioselectivity (Funhoff et al., 2006).

It is worth noting that several yeasts can assimilate alkanes and, at least in some environments, they may have an important role in the biodegradation of alkanes in oil-contaminated sites (Schmitz et al., 2000). In those cases studied, the enzymes involved in the initial oxidation of the alkane molecule are microsomal cytochromes P450 (Zimmer et al., 1996; Ohkuma et al., 1998; Iida et al., 2000).

2.3. Alkane hydroxylases for long-chain *n*-alkanes

Several bacterial strains can assimilate alkanes larger than C_{20}. These strains usually contain several alkane hydroxylases. Those active on C_{10}–C_{20} alkanes are usually related to *P. putida* GPo1 AlkB or to *Acinetobacter* sp. EB104 cytochrome P450. However, the enzymes that oxidize alkanes larger than C_{20} seem to be totally different. For example, *Acinetobacter* sp. M1, which can grow on C_{13}–C_{44} alkanes, contains a soluble, Cu^{2+}-dependent alkane hydroxylase that is active on C_{10}–C_{30} alkanes; it has been proposed to be a dioxygenase that generates *n*-alkyl hydroperoxides to render the corresponding aldehydes (Maeng et al., 1996; Tani et al., 2001). A different *Acinetobacter* strain, DSM 17874, has been found to contain a flavin-binding monooxygenase, named AlmA, which oxidizes C_{20} to > C_{32} alkanes (Throne-Holst et al., 2007). Genes homologous to *almA* have been identified in several other long-chain *n*-alkane-degrading strains, including *Acinetobacter* sp. M1 and *A. borkumensis* SK2. A different long-chain alkane hydroxylase, named LadA, has been characterized in *Geobacillus thermodenitrificans* NG80-2 (Feng et al., 2007). It oxidizes C_{15}–C_{36} alkanes, generating primary alcohols. Its crystal structure has shown that it is a two-component flavin-dependent oxygenase belonging to the bacterial luciferase family of proteins (Li et al., 2008). Several bacterial strains can degrade > C_{20} alkanes using enzyme systems that have still not been characterized and that may include new proteins unrelated to those currently known.

3. The role of the rubredoxin-rubredoxin reductase system

The rubredoxin that transfers electrons to the AlkB active site is a small redox-active iron-sulfur protein. The AlkG rubredoxin of *P. putida* GPo1 is unusual in that it contains two rubredoxin domains, AlkG1 and AlkG2, connected by a linker, while rubredoxins from other microorganisms have only one of these domains. Several rubredoxins from alkane-degrading bacteria have been cloned and analysed in complementation assays for their ability to substitute *P. putida* GPo1 AlkG. They clustered into two groups. AlkG1-type rubredoxins cannot transfer electrons to the alkane hydroxylase, while AlkG2-type enzymes can do so and can substitute for AlkG (van Beilen *et al.*, 2002). AlkG1-type rubredoxins probably have other as yet unknown roles. In fact, rubredoxin/rubredoxin reductase systems are present in organisms that are unable to degrade alkanes, where they serve other functions. For example, they participate in oxidative stress responses in anaerobic bacteria, transferring reducing equivalents from NADH to superoxide reductases, or to rubredoxin : oxygen oxidoreductases, thereby reducing oxygen or reactive oxygen species (Frazao *et al.*, 2000).

4. Metabolism of the alcohols and aldehydes derived from the oxidation of alkanes

The terminal oxidation of alkanes by alkane hydroxylases generates primary fatty alcohols, which are further oxidized to aldehydes by an alcohol dehydrogenase (ADH). There are several kinds of ADHs. Some use $NAD(P)^+$ as electron acceptor, while others transfer electrons to cytochromes or to ubiquinone. Most $NAD(P)^+$-independent ADHs contain pyrroloquinoline quinone as prosthetic group, and are named quinoprotein ADHs. Some bacterial species contain several different ADHs that can be used for the assimilation of distinct alcohols. For example, *P. butanovora* has at least four different ADHs with different specificities towards primary and secondary alcohols (Vangnai and Arp, 2001; Vangnai et al., 2002). *Acinetobacter calcoaceticus* HO1-N contains at least two ADHs; one shows preference for decanol while the other one has higher activity towards tetradecanol. Genes coding for alcohol and aldehyde dehydrogenases are also present in the *P. putida* GPo1 OCT plasmid. However, they are not essential for growth on alkanes because of the presence of similar enzymes in the *P. putida* GPo1 chromosome (van Beilen et al., 1992; 1994). The secondary alcohols generated by subterminal oxidation of alkanes are oxidized to ketones by ADHs.

Gordonia sp. strain TY-5, which can grow at the expense of propane or of C_{13}–C_{22} alkanes, metabolizes propane via 2-propanol and contains three NAD^+-dependent secondary ADHs, all of which are expressed in propane-grown cells (Kotani et al., 2003). NAD^+-dependent secondary ADHs have been identified in other bacteria such as *Rhodococcus rhodochrous* PNKb1 (Ashraf and Murrell, 1990).

Recently, an aldehyde dehydrogenase (ALDH) involved in alkane degradation by crude oil-degrading Geobacillus thermodenitrificans NG80-2 was characterized in vitro (Li et al., 2010). The ALDH was expressed heterologously in *Escherichia coli* and purified as a His-tagged homotetrameric protein with a subunit of 57 kDa based on SDS-PAGE and Native-

PAGE analysis. The purified ALDH-oxidized alkyl aldehydes ranging from formaldehyde (C_1) to eicosanoic aldehyde (C_{20}) with the highest activity on C_1. It also oxidized several aromatic aldehydes including benzaldehyde, phenylacetaldehyde, o-chloro-benzaldehyde and o-phthalaldehyde. The ALDH uses only NAD(+) as the cofactor, and has no reductive activity on acetate or hexadecanoic acid. Therefore, it is an irreversible NAD(+)-dependent aldehyde dehydrogenase.

6. Regulation of genes involved in aerobic *n*-alkane utilization in bacteria

Early studies by Thijsse and Linden (1958) have shown that *n*-alkane utilization by Pseudomonas aeruginosa is inducible, and later the same trait has been demonstrated for other *n*-alkane-oxidizing bacterial species (Perry and Scheld, 1968). However, some uncertainty remains regarding the nature of the inducers, which seem to be species-specific. Utilization of *n*-alkanes has been shown to be induced by both alkanes and alkanols in *P. putida* and *Burkholderia cepacia* (Grund et al. 1975; Marin et al. 2001), while in *Pseudomonas butanovora* only alkanols and aldehydes have been identified as inducers (Sayavedra-Soto et al. 2005). The initial studies on the genetics of regulation of the *n*-alkane utilization have been performed on *P. putida* GPo1, described to utilize *n*-alkanes from C_5 to C_{14} (van Beilen et al. 1994) (Fig. 6)

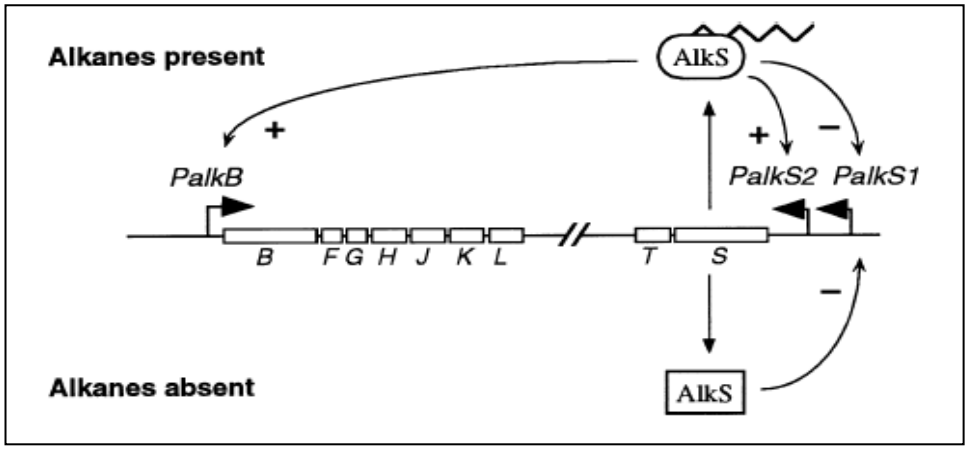

Figure 6. Regulation of the alkane oxidation pathway. Expression of the pathway is controlled by the AlkS regulator. In the absence of alkanes, very low levels of AlkS are produced from the sS-dependent PalkS1 promoter. In stationary phase, the activity of PalkS1 increases because of the increased levels of sS-RNAP, but transcription is kept low because of downregulation by AlkS. In the absence of alkanes, promoter PalkS2 remains inactive. When alkanes enter the cell, AlkS activates the PalkB and PalkS2 promoters, still repressing PalkS1. This leads to a rapid increase in AlkS levels that should allow a fast induction of the pathway. The simultaneous presence of a preferred carbon source modulates, or even represses, activation of the PalkB and PalkS2 promoters by AlkS.

The expression of the alkBFGHJKL genes involved in *n*-alkane utilization upon induction by n-alkanes have been shown to significantly affect the physiology of *P. putida* GPo1, leading to lower growth rate and filamentation on glucose as a carbon source (Chen et al. 1996). Authors have suggested that overproduction of the AlkB hydroxylase may be responsible for the observed phenotype. Furthermore, it has been shown that prolonged incubation of *P. putida* GPo1 on *n*-alkane-containing medium resulted in the loss of *n*-alkane-oxidizing activity, which was suggested to be due to the down-regulation of the alkBFGHJKL operon (Chen et al. 1996). This phenomenon can potentially cause a problem for a long-term biotechnological process based on *n*-alkane utilization/degradation because it would rely on an enzyme system required for good productivity while at the same time being deleterious to the host if it is being overexpressed. To better understand the regulation of the alkBFGHJKL genes expression, Yuste et al. (1998) constructed a reporter *P. putida* strain carrying the lacZ reporter gene under control of the PalkB promoter, which initiates the expression of the *alk* operon. It has been demonstrated that the expression from the PalkB promoter depends strongly on the carbon source available to *P. putida*, implying the involvement of a catabolite repression in regulation of the *n*-alkane degradation. Further work on the molecular mechanism of *alk* operon regulation in *P. putida* has demonstrated that the positive regulator AlkS is expressed poorly during the exponential phase of growth, while its expression increases considerably when cells enter the stationary phase (Canosa et al. 1999). The subsequent study has revealed that AlkS expression is controlled by two promoters, and that this protein regulates its own expression both positively and negatively (Canosa et al. 2000). In the absence of *n*-alkanes, expression of the alkS gene occurred mostly from the σS-dependent PalkS1 promoter, which provides for relatively high level of expression only during stationary phase. When the *n*-alkanes have been present in the growth medium, AlkS was shown to strongly repress the PalkS1 promoter, while activating its own expression

from a second promoter, PalkS2. Transcription from PalkS2 has been shown to be a subject for catabolite repression, and it has been suggested that the expression of *alkS* is regulated by a positive feedback mechanism. Such mechanism should allow both rapid induction of the *n*-alkane utilization pathway, and a fast downregulation thereof when the n-alkanes are consumed.

Regulation of *n*-alkane utilization has been studied in some detail for *A. baylyi* ADP1, which is capable of degrading long-chain *n*-alkanes with carbon chain length of C_{12} and more. Studies by Ratajczak et al. (1998) revealed that expression of alkane hydroxylase *alkM* in ADP1 is induced by C_7 to C_{11}, which are not utilized by this strain, and long-chain *n*-alkanes from C_{12} to C_{18}, which are metabolized. Transcription of *alkM* depended strictly on the AlkR, positive regulator of the AraC-XylS family normally expressed at a low level. The mechanism of regulation of *n*-alkane degradation by AlkR in *A. baylyi* ADP1 must therefore be different from the one involving AlkS in *P. putida* GPo1. While AlkS induces the expression of the *P. putida* GPo1 alkBFGHJKL operon in response to *n*-alkanes and the respective primary alcohols (Grund et al. 1975), the ability of AlkR to activate *alkM* expression in *Acinetobacter* depends on the presence of *n*-alkanes with chain lengths above C_6 (Ratajczak et al. 1998). Moreover, an inhibitory effect of oxidized *n*-alkane derivatives on the expression of *alkM* in ADP1 has been noticed. It seems interesting that the genes for rubredoxin, and rubredoxin reductase are constitutively expressed in ADP1 and do not seem to be subject to regulation in response to the presence of *n*-alkanes (Geißdörfer et al. 1999). It seems plausible that both these proteins, in addition to alkane utilization, are involved in other redox reactions important for normal cell functioning. Constitutive expression of rubredoxin and rubredoxin reductase genes have also been demonstrated in *P. aeruginosa* (Marin et al. 2003) and *Acinetobacter sp.* M-1 (Tani et al. 2001). In the latter strain, capable of degrading *n*-alkanes with chain length up to C_{44}, the expression of two alkane hydroxylases, AlkMa and AlkMb,

has been shown to be modulated by two independent regulatory proteins, AlkRa and AlkRb, respectively. It has been demonstrated that, while *alkMa* expression was induced by long-chain *n*-alkanes >C22, *alkMb* expression was preferentially induced by the *n*-alkanes with chain lengths of C_{16} to C_{22}. A regulation strategy different to that reported for the *P. putida* GPo1 *alk* genes was also found in *A. borkumensis*. In strain AP1, expression of the two genes encoding pAHs (*alkB1*, *alkB2*) was induced by *n*-alkanes, but transcription levels were found to be significantly different, possibly due to binding of transcriptional activator AlkS only to promoter PalkB1 of the *alkB1* gene. Unlike in GPo1, the expression of the *alkS* gene was not induced by *n*-alkanes, and an AlkS binding site was not detected upstream of alkS (van Beilen et al. 2004).

Aims

The aims of this work were to investigate the alkane degradation system of *Gordonia* SoCg by cloning and sequencing the *alk* locus; to analyze the gene expression in relation to the time course of *n*- alkanes consumption, and to identify the metabolic intermediate of *n*-hexadecane hydroxylation. Functional expression in heterologous hosts, unable to use *n*-alkanes, and the *Gordonia alkB* disruption mutant were performed in order to confirm the role of the AlkB in degradation of *n*-alkanes longer than C_{16}.

Materials and Methods

6.1 Bacterial strains, culture conditions and vectors.

Gordonia sp. strain *SoCg*, isolated from a hydrocarbon-contaminated Mediterranean shoreline on a mixture of hydrocarbons, was used in this work (Quatrini et al., 2008). It grows on the mineral medium Bushnell-Haas (BH, Difco) with a wide range of *n*-alkanes as a sole carbon source, from C_{12} up to C_{36} but does not grow on short chain *n*-alkanes.

Gordonia was routinely grown on JM medium (Puglia, et al., 1995) or on liquid mineral Bushnell Haas Medium (BH, Difco) supplemented with 10 mM *n*-alkanes directly supplied in the liquid medium (*n*-hexadecane, C_{16}) or supplied as finely ground powder (*n*-triacontane, C_{30}). In biodegradation kinetic experiments *n*-alkanes were added to BH medium as an *n*-hexane solution, once established that *n*-hexane is not toxic nor utilized by the strain. In solid cultures on BH agar, *n*-hexadecane was supplemented as vapor as described elsewhere (Quatrini et al. 2008).

Bacterial strains, commercial cloning vectors and plasmids constructed in this study are described in Table 1. *E. coli* was routinely grown in Luria medium (Sambrook, et al., 1989) and *S. coelicolor* in JM medium (Puglia et al., 1995).

Plasmid and chromosomal DNA purification, enzymatic digests, ligations and bacterial transformations were performed using standard molecular techniques (Kieser et al., 2000) or according to the manufacturer's instructions. All primers used for PCR amplification were synthesized by Invitrogen and are listed in Table 3. The 16S gene was amplified using primers rD1 and fD1 (Weisburg et al., 1991) in a 20 μl volume reaction containing 1 μl of chromosomal DNA, 0.2 μM of each primer, 0.2 mM of dNTPs and 1.5 Units of recombinant Taq DNA Polymerase (Invitrogen, Life Technologies). PCR was carried out in a Biometra Thermocycler using the following programme: 94°C for 5 min; 30 cycles of 1 min at 94°C, 1 min 55°C and 2 min at 72°C; final extension 72°C for 7 min. Where it is not differently

specified, the PCR was carried out under the same conditions but with an annealing temperature of 62°C.

6.2 PFGE analysis.

Gordonia SoCg was grown in 25 ml of YEME (Kieser et al.,2000) with 34% sucrose, 5mM $MgCl_2$ and 0,5% glycine for 40h at 30°C. Cells were harvested by centrifugation at 4,000 x *g* for 30 min, washed with 25mL TE25-sucrose (25mM Tris-HCl pH8, 25mM EDTA pH8, 0,3M sucrose) and centrifuged again. Then they were resuspended in TE25 sucrose to give OD_{600nm}= 1.9-2.0 and mixed with 1.5% low-melting-point SeaPlaque agarose (FMC, Rockland, Maine) in TBE (0.49 M Tris, 0.49 M boric acid, 0.001 M EDTA; pH 8). The resulting mixture was pipetted into plug molds (Bio-Rad Laboratories, Richmond, Calif.) and treated according to Kieser et al., 2000. After solidification at 4°C for 15 min, the agarose plugs were pushed out of the molds into 5 ml of a lysozyme solution (1 mg/ml) in TE25-sucrose. After incubation at 37°C for 2 h with swaying, the plugs were transferred to 5 ml of NDS (0.5 M EDTA, 0.01 M Tris [pH 9], 1% [wt/vol] lauroyl sarcosine) containing proteinase K (1 g/ml) and incubated at 50°C for 20 to 40 h. The proteinase K solution was pre-digested at 50°C for 1 h. For pulsed-field gel electrophoresis (PFGE), the plugs were inserted into the wells of 1% SeaPlaque agarose gel in TBE. Electrophoresis was performed with TBE 0.5X as the running buffer at 14°C by using a CHEF DRII PFGE system (Bio-Rad), at 160V, 400mA, n/s20 sec., e/w 20 sec., for 18h.

6.3 Southern hybridization analysis.

Undigested DNA separated by PFGE was transferred to Hybond N nylon membranes (Amersham International plc, Buckinghamshire, United Kingdom), according to the protocol for large DNA fragment transfer (Lee et al., 1991). Southern hybridization was carried out

by using the protocols provided by Amersham. To verify *alkB* localization in *Gordonia SoCg*, an *alkB* probe (AH), was obtained by digestion of plasmid pGAH (Table 1), and labeled with the digoxygenin (DIG) system (Boehringer Manheim Biochemicals, Indianapolis, Ind.). To verify the correct insertion in the *S. colelicor* M145 genome of the *pIJalKB* integrative vector, two different probes were used: *alkB* entire gene, PCR amplifyed from *Gordonia SoCg* genomic DNA and a 750 bp internal region of apramycin gene (Table 2), PCR amplified from pIJ773 plasmid. In order to verify *alkB* gene distruption, DIG-labelled alkCg341 was used in Southern analysis, that was carried out as previously described.

6.4 Cloning and sequence analysis of *Gordonia SoCg alk* genes.

Chromosomal DNA was isolated from *Gordonia SoCg* according to the method of Kieser et al., (2000). To clone the *alk* genes, suitable restriction fragments were identified in chromosomal DNA digested with *BamHI* and *BglII* enzymes, by Southern analysis (see above). Fragments in the range of 8 to 10 kb were cut out from a preparative agarose gel, isolated and ligated between the appropriate sites of *BamHI*-digested and dephosphorylated pUC18. The ligation mixture was used to transform *E.coli* DH10B (Invitrogen) by electroporation. *E.coli* trasformants were selected on ampicillin (200µg/ml). The trasformants were identified by colony PCR, using the couple of primers AH+ (Quatrini et al., 2008; Table 2). Plasmid pUC::*alk* cluster (p*alk*68) (Table 1) was isolated using the Miniprep Column Kit (Qiagen); its size was estimated by RFLP analysis and a 4.2kb region of the insert was sequenced by using the primers in Table 2.

Alternative monoxigenases was searched using different primers pairs (Table 2) without significant results.

6.5 Heterologous expression of *alkB* in *S. coelicolor* and *E. coli*

The *alkB* gene was amplified using primers alkBamHIfor and alkBNdeIrev (primers table) from *Gordonia SoCg* genomic DNA and cloned in pGEM-T-easy vector (Promega). The resulting pGalkB (Table 1) was sequenced and digested BamHI and NdeI to follow ligation in frame with tipAp of pIJ8600. The pIJalkB derivative (Table 1) were isolated and then transformed into *E.coli* ET12567 by electroporation. The ETpIJalkB was used to transform *S.coelicolor* M145 by coniugation according to *Kieser et al*., 2000. The resulting exconiugants were selected on apramycin. To evaluate the correct *alkB* gene expression in *S.coelicolor* M145, total RNA was extracted after thiostreptone induction in rich medium at 30°C, and used in RT-PCR assay as described below.

The *alkB* gene was also cloned in pRSETB expression vector (Invitrogen), by PstI-NcoI site from pGalkB, to an in frame ligation with T7 RNApol. The recombinant plasmids were transformed into *E. coli* BL21 (DE3)pLysS (Invitrogen) and the transformants were selected on LB supplemented with ampicillin (200 mg/ml). Alkaline Phosphatase-immunoconiugated antibody was used to verify the *alkB* gene correct translation in *E.coli BLpRalKB.* Crude extracts were collected after incubation in LB medium from *E.coli* BLSETB and *E.coli* BLpRalkB with IPTG as inducer and also from *E.coli* BLpRalkB without inducer, using manual conditions (Ni-NTA purification system, Invitrogen). About 0,500 µg of each soluble and insoluble protein fraction were loaded in SDS-PAGE at 20mA, 150V for 50min and transfer to Hybond membrane for Western Blot analysis, using Towbin buffer at 150V, 20mA, 1h.

6.6 Time course of growth on *n*-alkanes.

Gordonia SoCg was grown in 30mL JM medium (Puglia et al., 1995) for 48h at 30°C in a buffled flasks (250mL). Cells were washed three times with BH medium and resuspended in BH mineral medium to give an OD_{600nm}= 1.0. Afterwards, 1mL (about 1mg dry weigth) of

bacterial suspension was inoculated in three different buffled flasks (1L) containing 120mL BH supplemented with 10mM *n*-hexane, *n*-triacontane and fructose, respectively. The flasks were incubated at 30°C. At 16, 18, 22, 24, 40, 42, 46, 48, 64, 66, 70, 72, 96 h, and 1mL aliquots were sampled from each buffled flask in triplicates, centrifuged at 4000 x *g* and the pellet dried at 65° C until it reached a constant weight. *E.coli* BLSETB and BLalkB was grown in 3mL LB medium (Invitrogen) for 12h at 37°C in a 50mL flask. Cells were washed three time with phosphate buffer pH 7.2 and resuspend in BH to give an OD_{600nm}= 1.0. Afterwards, 300 µl of bacterial suspension was inoculated in two different 500mL flask containing 100mL BH supplemented with 10mM *n*-dodecane, *n*-hexane and fructose, respectively. The flasks were incubated at 37°C. At 5, 15, 30, 45 h 1mL aliquots were sampled from each buffled flask in triplicates and analysed at OD_{600nm}, using BH with *n*-alkane respective as auto zero.

6.7 Time course of *n*-alkane consumption.

To determine *Gordonia SoCG n*-alkane utilization, a rich 48h pre-inoculum in 30mL JM medium (Puglia et al., 1995) was washed and suspended in BH broth to give a final OD_{600nm} = 1.0 and then inoculated in 100mL glass tubes with 10ml BH medium supplemented with 10mM of each *n*-alkane, at the same concentration used during growth curves construction. A total of 48 tubes that were incubated at 30°C under shaking conditions.: twelve tubes for each *n*-alkane were inoculated with *Gordonia SoCg*, and twelve were left not inoculated to be used as abiotic controls. Residual long-chain *n*-alkanes were extracted, using *n*-hexane and sonication, from the whole tube content in triplicates after 22, 46, 62 and 96h of incubation, and analysed with GC-MS as previously described (Quatrini et al., 2008).

6.8 Analysis of the metabolic intermediates from *n*-alkanes oxidation pathway.

The metabolic intermediates resulting from incubation on C_{16} and C_{30} of SoCg, M145-AH and BL21-AH expressing the Gordonia *alkB*, were analyzed by solid-phase micro extraction (SPME) coupled with GC-MS. *Gordonia* SoCg, *Gordonia* SoCg Ω*alkB, S.coelicolor* M145-AH and *S.coelicolor* M145 carrying the empty pIJ8600 were grown in JM medium, washed and re-suspended in BH medium to a final OD_{600} of 1.0 as previously described. *E.coli* BL21-AH and *E.coli* BL21 carrying the empty pRSET-B were grown O.N. in LB medium, washed three times with phosphate buffer pH 7.2 and suspended in the same volume. One mL of cells was inoculated in 100mL glass tubes with 10mM of each *n*-alkane in the presence of inducers (IPTG in *E. coli* tubes according to the Ni-NTA purification system instructions by Invitrogen and Thiostrepton 10ng mL^{-1} in *Streptomyces* tubes) and incubated at 37°C for 6h with shaking. Abiotic controls were incubated under the same conditions and analyzed in parallel.

The entire suspensions were analyzed by immersing the SPME fiber coated with 85 µm polyacrylate (PA) and equipped with holder for manual injection. The time needed to reach equilibrium between the amount of analyte adsorbed by the polymeric film and the initial concentration of the analyte in the sample matrix, during the SPME sampling, is dependent on the properties of both the analyte and the matrix (de Pasquale et al., 2007) and in our study was 20 min at 45°C. Prior to use, the fiber was conditioned at 300 °C for 2 h, each in the GC injector port. HP-5MS 5% Phenyl Methyl Siloxane Capillary column was used to perform the gas-chromatographic separations. The oven initial temperature was 80 °C with an helium constant flow, corresponding to the nominal head pressure of 9.37 psi. The temperature increased of 5 °C min^{-1} to 280 °C, temperature held for 20 minutes. The ionization spectra were obtained as mentioned above. The analytical identification and quantifications were carried out using standard grade compounds purchased from Sigma-Aldrich (USA) and commercial NIST 2005 mass spectra library search database.

6.9 Total RNA isolation, RT-PCR analysis, and absolute qRT-PCR.

The cells were broken by using 3 mg of lysozyme/ml in P-buffer (Kieser et al., 2000) and total RNA was isolated using the RNeasy midi-kit (QIAGEN). DNaseI (Roche) treatment was performed at 37°C for 1 h, and ethanol precipitation in the presence of 0.1 vol 3 M sodium acetate allowed recovery of the DNase treated total RNA. After a washing step with 70% ethanol and air drying, the RNA pellet was resuspended in 50µl of water. Reverse transcription-PCR (RT-PCR) was performed by using the Superscript One-Step RT-PCR kit (Invitrogen) with about 0.1 µg of total RNA as a template, primer pairs internal to *alk* genes (AHqRT) (Table 3) and the conditions indicated by the supplier, routinely using 35 PCR cycles. For each reaction, a negative control with *Taq* polymerase and without reverse transcriptase was included, in order to exclude DNA contamination. Expression was quantitatively analyzed by real-time RT-PCR using the Applied Biosystems 7300 real-time PCR system (Applied Biosystems). A high capacity cDNA archive kit (Applied Biosystems) was used, according to the manufacturer's instructions, to retrotranscribe 5 µg of total RNA DNA free, extracted after 22 h of growth in mineral salt medium with hexadecane, triacontane and fructose respectively. Then, 3 µl of the cDNA was mixed with 10 µl of SYBR green PCR master mix (Applied Biosystem) and 5 pmol of each primer (Table 3) in a final volume of 20 µl. The PCR was performed under the following conditions: 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and1 min at 62°C. Eventually, a dissociation reaction was performed with the following conditions: a 1-min step with a temperature gradient increase of 1°C per step from 55 to 99°C. This last reaction allowed the melting curve of the PCR products and, consequently, determine their specificity. A negative control (distilled water) was included in all real-time PCR assays, and each experiment was performed in triplicate. Standards for the *alkB* genes were constructed from

p*alk*68, purified by Miniprep Column Kit (Quiagen) and measured spectrophotometrically at 260 nm.

This was diluted in a 10-fold series to create the standards for a six-point standard curve (0,5 $x10^2$ to 0,5 x 10^6 molecules) that was run in duplicate with each set of samples. The number of copies per microliter was calculated as follows:

– Molecular weight of *palk68* (standard template) = 11433 bp x 660 Da = 7,54x 10^6 g

– 1 molecule or 1 copy of fragment = 7,54x 10^6/6.02 x 10^{23} = 1,25x 10^{-17} g

– therefore 10 ng of template contain $10x10^{-9}$/1,25 x 10^{-23} copies = $8x10^{14}$ molecules

6.10 *Gordonia SoCg* electrocompetents cells

About 2mg di *Gordonia SoCg* were inoculated 3 days in 25mL YEME with 2g/L of Glycine at 30°C. The reach biomass was collect at the bottom of the 50mL tube and pretreated for 15 min in a ultrasonic batch. Pellet was obtained after centrifugation at 5500xg 30 min and rapidly incubated with 10 g/L chilled glycerol on ice for 1h. Following this step, cells were washed three times with 10 g/L chiled glycerol and finally, resuspended in 3 mL of 10 g/L glycerol, aliquoted into 200 µl samples and stored at -80°C.

6.11 Construction of *Gordonia* SoCg *alkB* disruption mutant

In order to obtain *Gordonia* SoCg electrocompetent cells 2mg (w.w.) cells were inoculated 3 days in 25mL YEME (9) with 2 g L^{-1} of Glycine at 30°C. The rich biomass was collected at the bottom of the 50mL tube and pretreated for 15 min in an ultrasonic batch. The pellet obtained after centrifugation at 5500 x *g* for 30 min was immediately incubated with 10 g L^{-1} chilled glycerol on ice for 1h. Cells were then washed three times with 10 g L^{-1} chilled glycerol and finally, re-suspended in 3 mL of 10 g L^{-1} glycerol, aliquoted into 200 µl samples and stored at -80°C.

The apramycin resistance cassette, including its own promoter and oriT, was extracted from plasmid pIJ773 by digestion with *EcoRI* and *HindIII* and cloned into pUC18. The correct recombinant plasmid was checked by sequencing and the cassette was excised using *EcoRI* and filled-in, using the Klenow fragment enzyme (Roche), to obtain blunt ends. The apramycin resistance cassette was inserted into the unique *Ale*I site within the *alkB* gene of palk68. The resulting plasmid (palkapra, Table 1) was *XbaI-NcoI* digested to obtain an *alkB*::apra linear fragment (Fig.1), that was introduced by electroporation into *Gordonia* SoCg electrocompetent cells. Apramycin-resistant transformants were selected on apramycin, and gene disruption by double-crossover homologous recombination, was confirmed by Southern analysis using the DIG-labeled alkCg341 as a probe.

Nucleotide sequence accession numbers

The *Gordonia* SoCg 16S rDNA sequence has been submitted to Genbank under accession no. AY496285.2 and the 4,472-bp *palk68* insert under accession no. HQ026811.

7. Results

7.1 Identification and properties of strain SoCg.

The *n*-alkane degrader *Gordonia* sp. SoCg was isolated from a hydrocarbon-contaminated Mediterranean shoreline. This strain is able to grow on *n*-alkanes of different length from dodecane (C_{12}) to hexatriacontane (C_{36}) as the sole C source (Quatrini et al., 2008) and unable to grow on *n*-octane, or shorter *n*-alkanes, although it is not inhibited by *n*-hexane. Complete 16S rDNA sequence analysis showed the highest similarity (98% identity, 1442/1470 nt) to the DNA sequence of *Gordonia amicalis* strain T3 (Genbank accession no EU427321.1), which is a tert-amyl methyl ether degrader, recently isolated from a hydrocarbon contaminated soil (Purswani, J. et al., 2008). Pulsed Field Gel Electrophoresis (PFGE) of undigested DNA revealed the presence of a criptic plasmid of the apparent size of 150kb, as determined by comparison with a linear DNA molecular weight marker (Fig.7). Using a 570 bp DIG-labelled *alkB* fragment, derived from plasmid pGAH, as a probe (alkCg23) in a Southern hybridization of PFGE-separated DNA, we localized the *alkB* gene on the chromosome of *Gordonia* SoCg (Fig.8).The same probe hybridized to only one band in the genomic DNA that had been digested with various restriction enzymes (Fig.9). This results confirms that strain SoCg harbors in its chromosome only one copy of *alkB* gene, as previously revealed by PCR (Quatrini et al., 2008).

As SoCg degrades a large range of long chain *n*-alkanes, we also tried to find in its genome other *alkB*-unrelated putative genes involved in long chain *n*-alkane degradation, using PCR. The primer couple designed on the consensus sequence of *Geobacillus thermodenitrificans* NG80-2 *ladA* and of other Gram-positive strains gave two different amplification products, whose sequence was unrelated to known alkane hydroxylase genes (data not shown). A second couple of primers CF and CR (Kubota et al., 2005), used to amplify the conserved region of the p450-CYP153 family genes gave no amplification product.

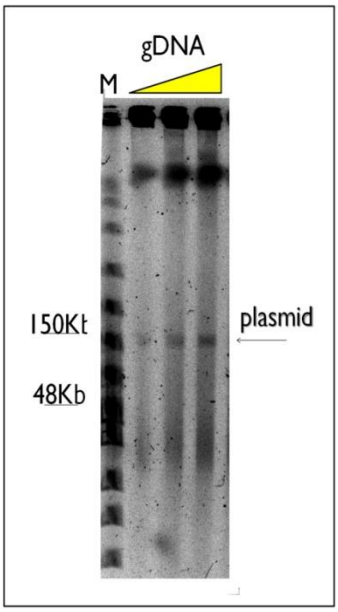

Figure 7. Plasmid identification.
PFGE analysis of (gDNA) genomic DNA of *Gordonia sp*. SoCg strain loaded at different concentration. M: molecular weight marker.

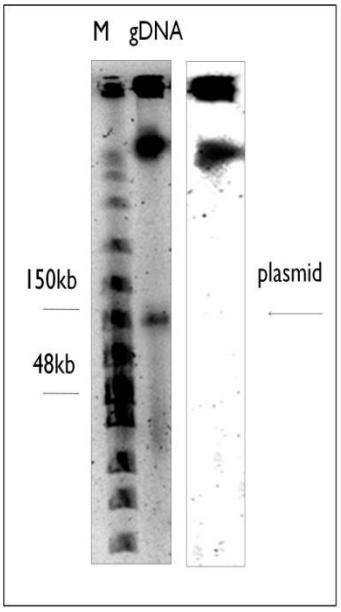

Figure 8. Localization of the *alkB* gene.
Southern blot of the (gDNA) genomic DNA, running in PFGE, with alkCg341 as probe (see Table 2 and Fig.20). M: molecular weight marker.

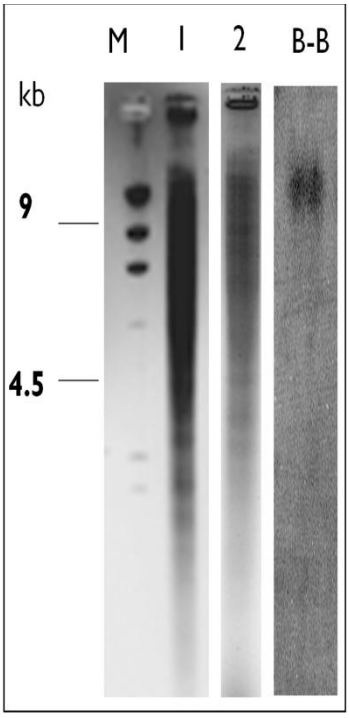

Figure 9. Estimation of copy number of the *alkB* gene. Agarose gel electrophoresi of gDNA digested with 1) BamHI, (2) BglII and Southern blot of the (gDNA) genomic DNA, with alkCg341 as probe (see Table 2 and Fig.20), digested with (B-B) BamHI and BglII. M: molecular weight marker.

7.2 Growth on long chain *n*-alkanes and biotrasformation kinetics.

Growth of *Gordonia* sp. SoCg in mineral broth supplemented with *n*-hexadecane and *n*-triacontane, respectively, as the sole carbon source, was followed for 96 h. An increasing biomass production was observed on both *n*-alkanes up to 46 h followed by a biomass decline (Fig.10). Comparison of the two curves reveals that *n*-hexadecane supports higher biomass accumulation (3.4 mg mL^{-1},) than *n*-triacontane (2 mg mL^{-1}), as measured after 46 h of growth. GC-MS analysis of residual *n*-alkanes showed that substrate consumption coincides with biomass production. Both *n*-alkanes, in fact, have almost completely disappeared within 62h (with 11.1% residue of *n*-hexadecane and 1.9% of *n*-triacontane) with a rapid decline in the first 22 h (Fig.11). Consumption of *n*-triacontane is more rapid than that of *n*-hexadecane, in contrast with lower biomass accumulation on the longest alkane. Although it seems that the alkane monooxygenase is able to bind triacontane and hexadecane with the same affinity, as demonstrated by their rapid consumption, its efficiency is gradually reduced with the increasing of chain length as already reported for other strains (Feng, et al., 2007; Sameshima et al., 2008; Smits et al., 2002; Throne-Holst et al., 2006).

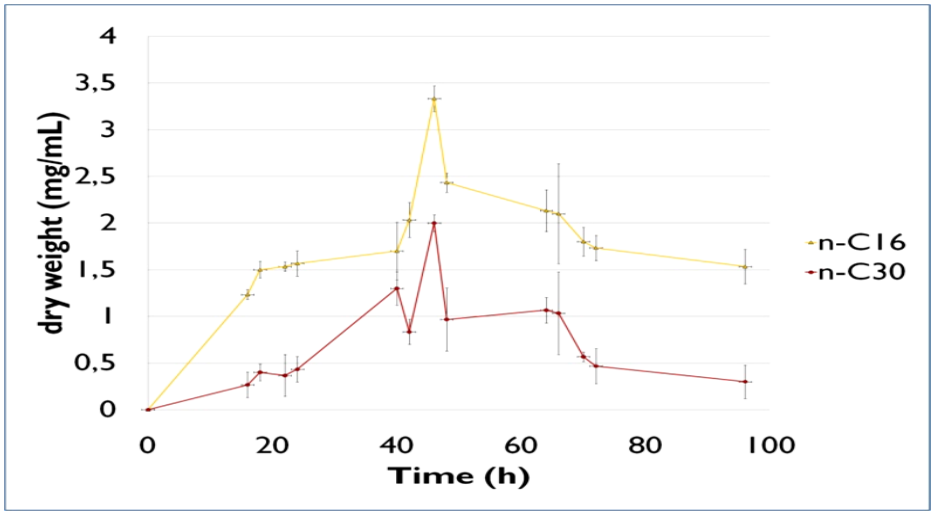

Figure 10. Growth on long chain *n*-alkanes. Time course of growth of *Gordonia* sp. SoCg wild-type on 10mM of (▲) *n*-hexadecane and (●) *n*-triacontane. Standard Error vertical bars are calculated from three independent determinations.

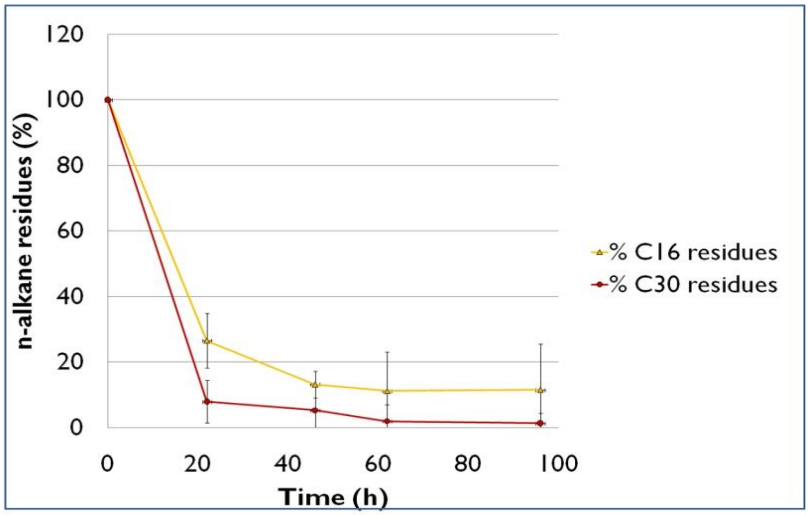

Figure 11. Degradation kinetics by *Gordonia* sp. SoCg.
Time course of consumption of (▲) *n*-hexadecane and (●) *n*-triacontane, determined by GC-MS and expressed as *n*-alkane residue in respect to uninoculated controls. Standard Error vertical bars are calculated from three independent determinations.

7.3 Cloning and sequence analysis of SoCg *alk* genes.

In order to isolate the unique alkane hydroxylase gene and its flanking regions, we cloned a *BamHI – BglII* fragment into pUC18 obtaining the plasmid palk68 (Fig.12). An internal fragment of palk68 was completely sequenced using chromosome walking strategy with primers listed in Table 2 and oriented as shown in Fig.17. This strategy allows to take a 4,472-bp nucleotide sequence that exhibited an overall identity of 79% with the AH encoding system of *Gordonia* sp. TF6 (Fuji et al., 2004) and of 78% with that of *Rhodococcus* Q15 and NRRL B-16531(Whyte et al., 2002). Sequence analysis revealed six consecutive open reading frames (ORFs) whose products showed the greatest amino acid sequence identities with the complete sequences of alkane hydroxylase components and of a putative regulatory proteins (Fig.13) (Table 3). The ORFs were designated AlkB (alkane 1-monoxygenase), RubA3 (rubredoxin), RubA4 (rubredoxin), RubB (rubredoxin reductase), AlkU (putative TetR-like regulator), according to other homology to know genes (Whyte et al., 2002; Fuji et al., 2004) (Table 3).

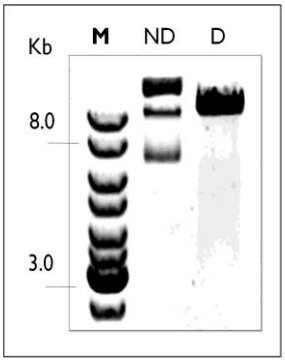

Figure 12. AleI-enzymatic digestion of palk68. 1% Agarose gel electrophoresis of (ND) undigested and (D) digested palk68. M: molecular weight marker

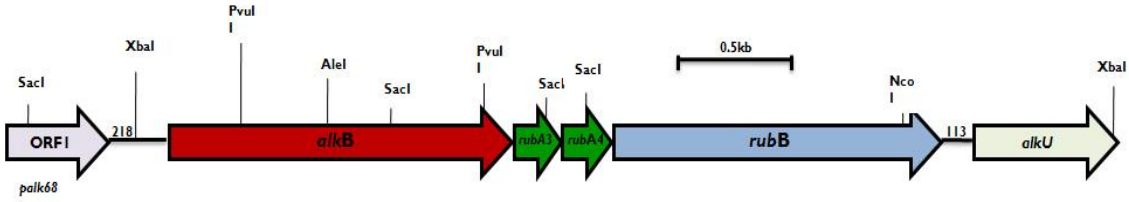

Figure 13. Schematic representation of the 4,472-bp region of *Gordonia* sp. strain SoCg. Genetic organization of the gene cluster and restriction map. The orientation of identified genes is indicated by arrows.

7.4 Characterization of the AH system of *Gordonia SoCg*

7.4.1. *alk* genes expression analysis

The expression of *alk* genes in the presence of long chain *n*-alkanes, was analyzed by reverse transcriptase-quantitative PCR. Total RNA was extracted from *Gordonia* SoCg cultures after 22h of growth at 30°C in mineral broth supplemented with *n*-hexadecane and *n*-triacontane as the sole carbon source, respectively. Total mRNA from fructose-grown cells was extracted and analysed as control (Fig.14). cDNA molecules were retrotranscribed from the total DNA-free RNA and used as templates to quantify *alkB*, *rubA3*, *rubA4*, *rubB* and *alkU* transcripts. Interestingly, the expression level of the four *alk* genes is generally higher on *n*-triacontane than on *n*-hexadecane, with *alkB* gene expression 2.5 fold higher on *n*-triacontane than on *n*-hexadecane (Fig.15).Our results show that both long chain *n*-alkanes induce the expression of all the *alk* genes but *alkU*. *alkU* has been found downstream of the *alkb-rubA-rubB* cluster in Gram-positive *n*-alkane degraders (Hamamura et al., 2001; Wentzel et al., 2007) and also in the genome of *Nocardia farcinica* (Ishikawa et al., 2004); it possesses helix-turn-helix DNA binding motifs and shows deduced amino acid similarity to putative regulatory proteins of the TetR family; however we found that it is not influenced by *n*-alkanes and further investigations are needed to assess its involvement in *n*-alkanes degradation.

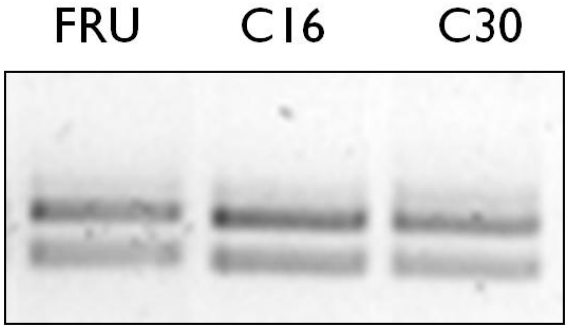

Figure 14. RNA isolation. 1% Agarose gel electrophoresis of total RNA extracted from *Gordonia sp*. strain SoCg after 22h of growth at 30°C in mineral broth supplemented with (FRU) fructose, (C16) *n*-hexadecane, (C30) *n*-triancontane, as sole carbon source.

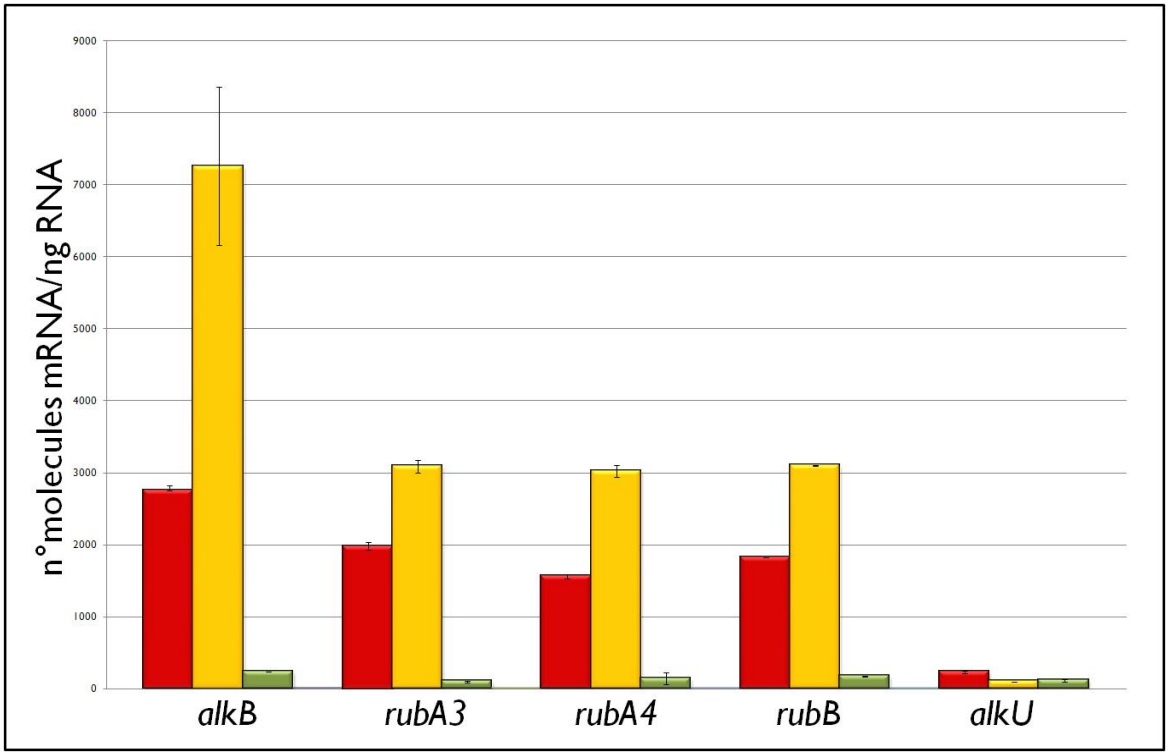

Figure 15. Absolute real-time RT-PCR analysis of *alk* genes of *Gordonia sp.* strain SoCg.
mRNA levels after 22h of incubation in the presence of (●) *n*-hexadecane, (●) *n*-triacontane and (●) fructose, respectively, where expressed as n° molecules /µg total RNA. Error bars are calculated from three independent determinations of mRNA abundance in each sample.

7.4.2. Long chain *n*-alkanes biotransformation

Metabolic intermediates were extracted by SPME from SoCg cells incubated with *n*-alkanes and identified by GC-MS. After 6h incubation in the presence of *n*-hexadecane as the sole C source, 1-hexadecanol was identified comparing its Kovats index and the electron impact mass spectra with those obtained by the injection of the authentic standards (Fig.16). When SoCg was grown on *n*-triacontane, on the contrary, we were unable to reveal the corresponding primary long-chain alcohol. Longer incubation time or other extraction methods (i.e. hexane extraction) did not lead to triacontanol detection (data not shown). Almost all alkane hydroxylase activities so far have been analyzed indirectly by mineralization of ^{14}C labeled *n*-alkanes (White et al., 2002) and growth assays (Smits et al 2002; Throne-

Holst et al., 2007). Only a few authors report the detection of metabolic intermediates of long chain *n*-alkanes metabolism in the naturally occurring *n*-alkane degrader (Whyte et al., 1998) or in a heterologous host (Fujii et al., 2004), or the activity of the purified protein. Here we report for the first time the biotransformation activity on hexadecane to the corresponding alcohol by a strain of *Gordonia*. Our results clearly indicate that *n*-hexadecane is metabolized via terminal oxidation pathway, like in other *n*-alkane degrading bacteria. As we did not detect any long chain alcohol on triacontane we can hypothesize that it is immediately used in the following reactions (Koma et al., 2001) or that it is undetectable because of its insolubility (Tani et al., 2001).

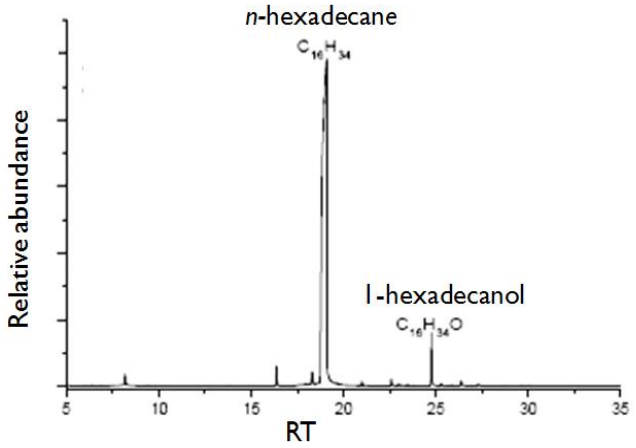

Figure 16. *n*-hexadecane bioconversion to 1-correspondent alcohol. GC-MS spectrum of solid phase microextraction (SPME) from cultures of *Gordonia sp.* strain SoCg after 22h of incubation in mineral medium supplemented with *n*-hexadecane.

7.4.3. The *alkB* disruption mutant

To investigate the rilevance of AlkB for degradation of long chain *n*-alkanes by *Gordonia* SoCg we inactivated the *alkB* gene by introducing an apramycin resistance cassette in a unique *AleI* restriction site that is present in position 576 of its nucleotide sequence (Fig.17). The recombinant strains were selected on apramycin solid medium and site specific apramycin cassette insertion was first analyzed by PCR amplification of internal fragments of the resistance cassette with primers apra750FR and apra750FR (Fig.18 A) (table 3), PCR amplification of recombinant fragment alkapra (Fig.18 B), PCR amplification of alkB (Fig.18 C) and of the region between the 5'-end of *alkB* and the 5'-end of *rubA3* (Fig.19).

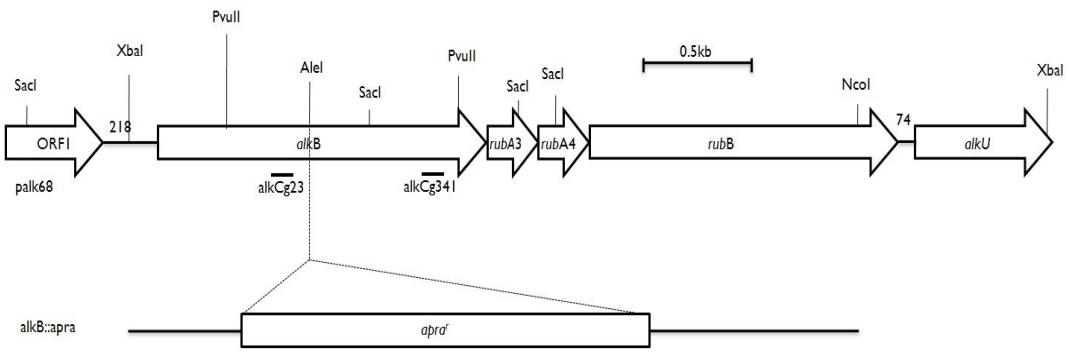

Figure 17. Schematic representation of the construction of alkB distruptant mutant. Genetic organization of the gene cluster and restriction map. The orientation of identified genes is indicated by arrows. Probes alkCg570 and alkCg341 were used in Southern hybridization for *alkB* localization and disruption mutant analysis, respectively.

Finally the *Pvu*II- digested genomic DNA of four positive clones was analyzed in a Southern hybridization experiment using the DIG-labeled alkCg341 probe (Fig.19). One strain, showing the correct insertion of the cassette, was named *Gordonia SoCg ΩalkB* and used for further experiments.

Strain *SoCg ΩalkB* exhibited poor growth on *n*-hexadecane (7 fold lower in respect to the wild type strain) (Fig.20) and, interestingly, no growth on triacontane (Fig.21), making it

evident that *alkB* disruption had a negative effect on solid *n*-alkanes metabolic pathway. When the disruption mutant was incubated in the presence of 1-triacontanol as the sole C source it was able to grow even better than the w.t. on triacontane (Fig.22).

On the other hand, *Gordonia SoCg ΩalkB* was still able to transform *n*-hexadecane into the corresponding alcohol (data not shown) suggesting that oxidation of *n*-hexadecane in the absence of *alkB* must be carried out by an unknown oxidation system that in any case does not allow efficient growth of the strain.

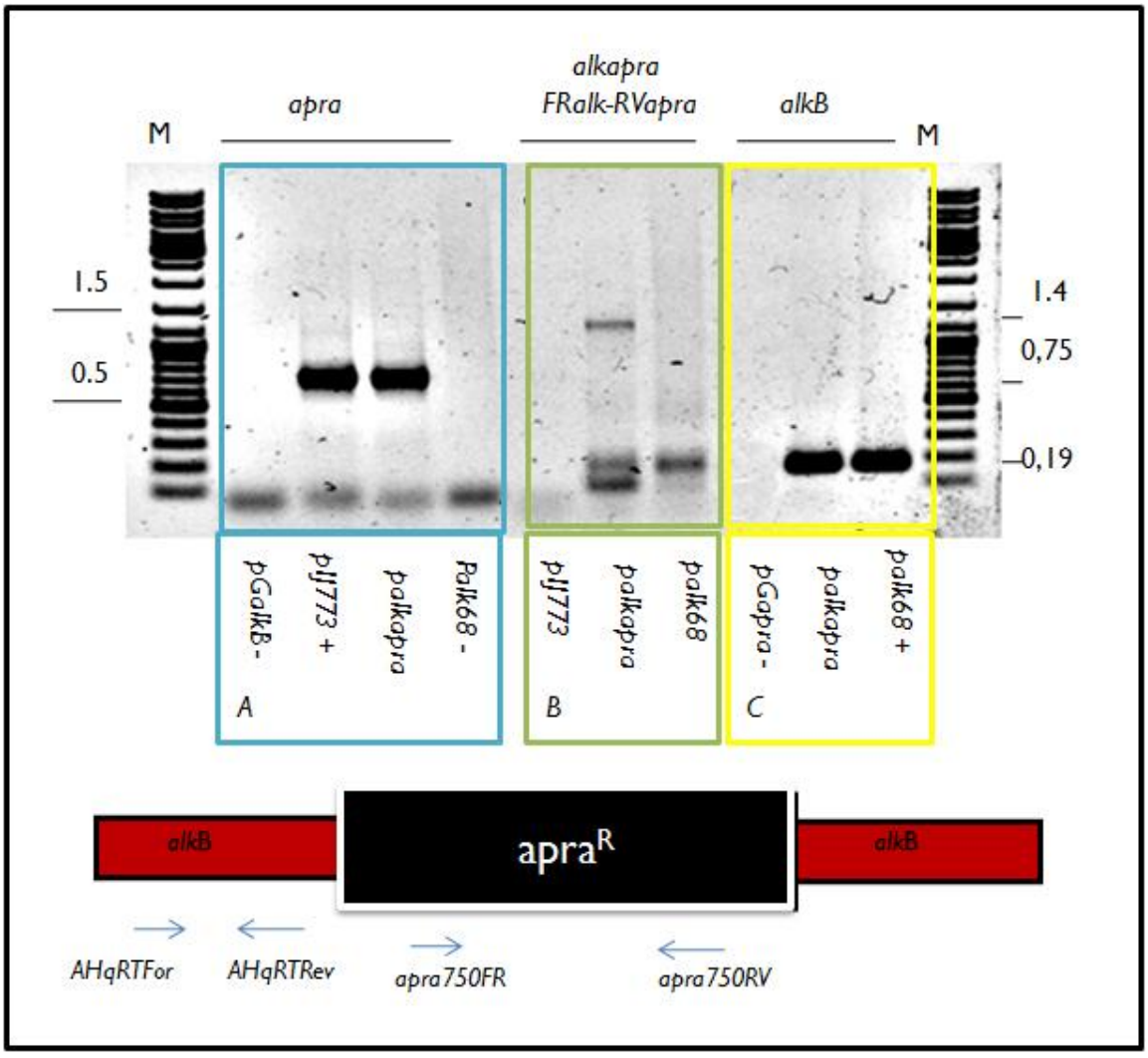

Figure 18. Schematic representation of the PCR-based selection of *E.coli* recombinant strains harboring distrupted *alkB* gene. 1% agarose gel electrophoresis of amplicons obtained with (A panel) apra750 primer pairs; (B panel) AHqRTFor and apra750RV; (C panel) AHqRT primer pairs. The correct amplification of apramycin internal region and *alkB* internal region indicate that apramycin resistance cassette was successfully introduced into palk68 and finally, the correct amplification of the chimeric construct alkapra indicate the exact insertion of cassette into unique AleI restriction site (All primer pairs are listed in Table 2).

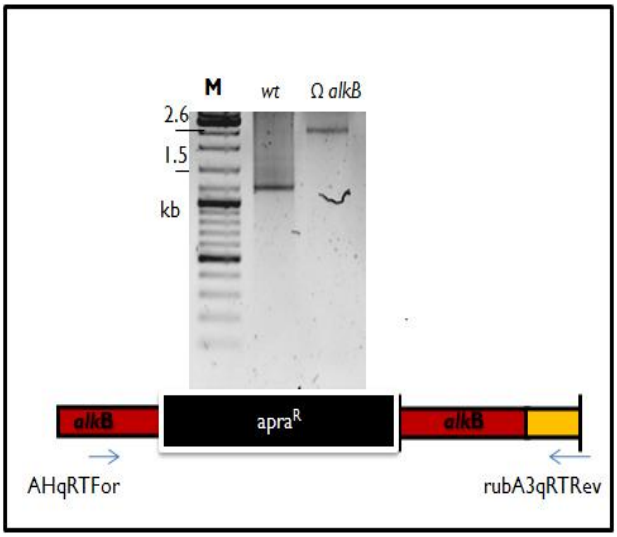

Figure 19. Schematic representation of PCR-based selection of *Gordonia sp. strain* SoCg recombinant strain harboring distrupted *alkB* gene. 1% agarose gel electrophoresis of amplicon obtained with AHqRTFor and rubA3qRTRev primers, using (wt) wild type genomic DNA and (Ω *alkB*) distruptant mutant genomic DNA. M: molecular weight marker.

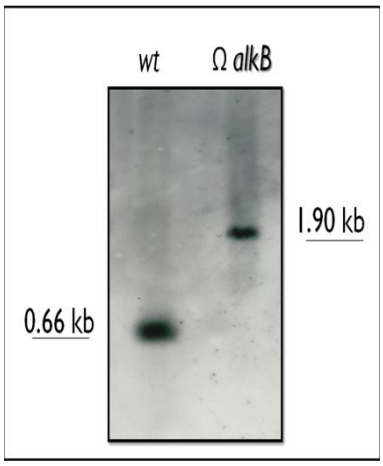

Figure 20. Southern analysis of the PvuII-digested genomic DNA extracted from the (wt) wild-type *Gordonia* sp. strain SoCg and (Ω alkB) the mutant strain *Gordonia* SoCg, with DIG-labeled probe alkCg341.

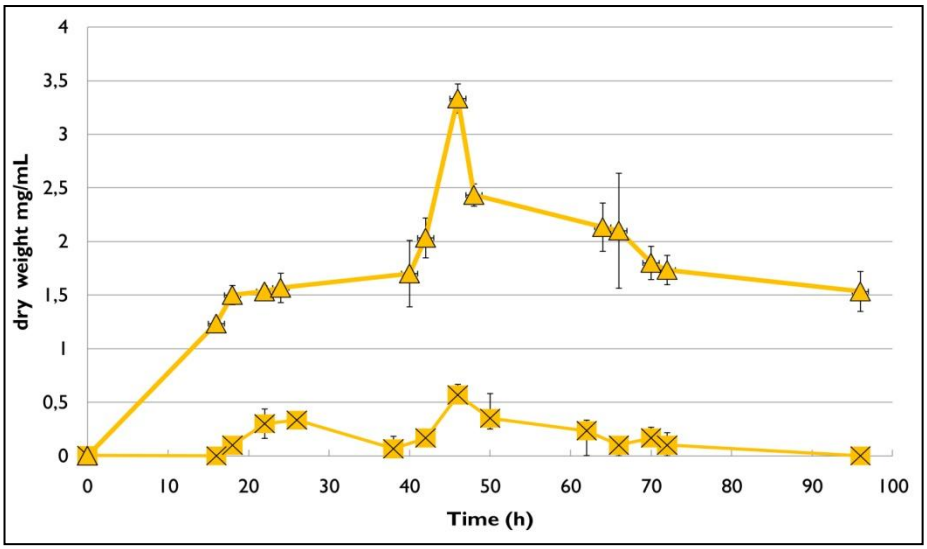

Figure 20. Comparison of biomass accumulation on *n*-hexadecane. Time course of growth of (▲) *Gordonia* sp. SoCg wild-type and (X) *Gordonia sp.* Ω alkB . Standard Error vertical bars are calculated from three independent determinations.

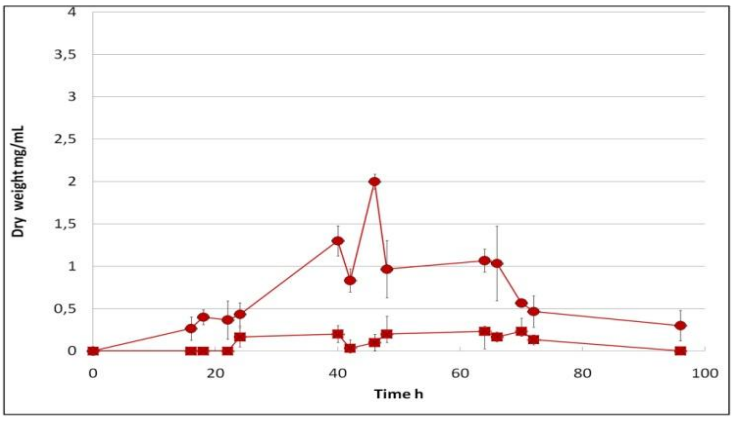

Figure 21. Comparison of biomass accumulation on *n*-triacontane. Time course of growth of (●) *Gordonia* sp. SoCg wild-type and (X) *Gordonia sp.* Ω alkB . Standard Error vertical bars are calculated from three independent determinations.

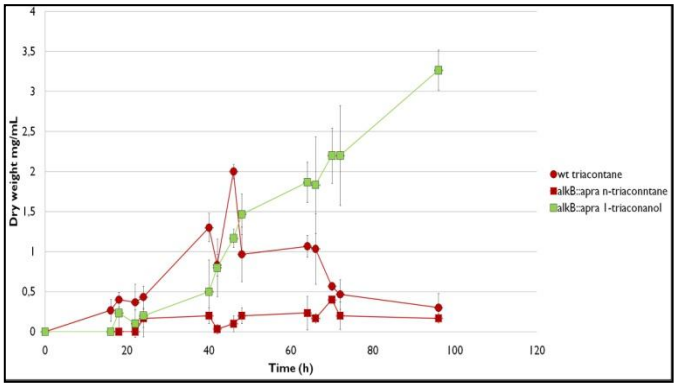

Figure 22. Comparison of biomass accumulation. Time course of growth of (●) *Gordonia* sp. SoCg wild-type and (X) *Gordonia sp.* Ω alkB on *n*-triacontane and ( ) on 1-triacontanol. Standard Error vertical bars are calculated from three independent determinations.

7.4.4. AlkB heterologous expression in *S. coelicolor* and *E. coli*

The unique *alkB* gene was heterologously expressed in *E. coli* BL21 using the expression vector pRSET-B. The expression of his-tagged AlkB in the resulting recombinant strain *E.coli* BL21-AH (Table 1) was confirmed by Western blot technique using alkaline phosphatase-conjugated anti-His tag monoclonal antibodies (Invitrogen) (Fig.23 A-B). The protein was detected mainly in the insoluble fraction of a crude extract of *E.coli* BL21-AH after 4 hours induction with IPTG (Fig.23 B). SPME/GC-MS analysis on bacterial cultures IPTG-induced for 4h revealed that *E.coli* BL21-AH was able to transform *n*-hexadecane into 1-hexadecanol (Fig.24). No hexadecanol or other products were detected using *E. coli* carrying pRSET-B (data not shown).

Although *alkB*-related alkane hydroxylase activity is known to be rubredoxin and NAD(P)H dependent (van Beilen et al., 2007), we obtained hexadecane hydroxylation in *E. coli* expressing only *alkB.* Fujii and colleagues defined *alkB, rubA3, rubA4* and *rubB* as the minimum component genes of the alkane hydroxylase systems. In fact these authors obtained biotransformation of n-octane to 1-octanol in E. coli TOP10 carrying plasmid pAL526 that contained the *Gordonia* TF6 *alk* cluster composed of the four genes. However the AlkB relative activity was not completely eliminated in the absence of the other alkane hydroxylase system components. Similarly two *E. coli* recombinants which expressed the *R. opacus* B-4 *alkB1* and *alkB2* genes were able to convert *n*-alkanes (C_5-C_{16}) to their corresponding alcohols in anhydrous organic solvents (Sameshima et al., 2008).

When *n*-triacontane was used as substrate of *E.coli* BL21-AH, we were unable to reveal the corresponding primary long chain alcohol using SPME/GC-MS analysis (data not shown).

As *E. coli* may not be the appropriate host we decided to use *S. coelicolor* M145 to express the *Gordonia alkB*. This strain does not grow on *n*-alkanes but recently *n*-hexadecane degrading *Streptomyces* species have been isolated (Barabas et al., 2001). Using the integrative

plasmid pIJ8600 (Fig.25) we obtained the recombinant strain *S.coelicolor* M145-AH (Fig.26) expressing *alkB* in the presence of thiostrepton (Fig.27) and is able to growth using *n*-hexadecane as the sole carbon source (Fig.28). SPME/GC-MS analysis showed the presence of 1-hexadecanol in *S.coelicolor* M145-AH cultures on *n*-hexadecane (Fig.29). This is the first study to achieve biotransformation of *n*-hexadecane to 1-hexadecanol using *S. coelicolor* expressing an alkane hydroxylase gene.

S.coelicolor M145 w.t. and *S. coelicolor* carrying pIJ8600 were unable to biotransform *n*-hexadecane (data not shown). However, using *S.coelicolor* M145-AH, we were still unable to reveal the corresponding primary long chain alcohol 1-triacontanol.

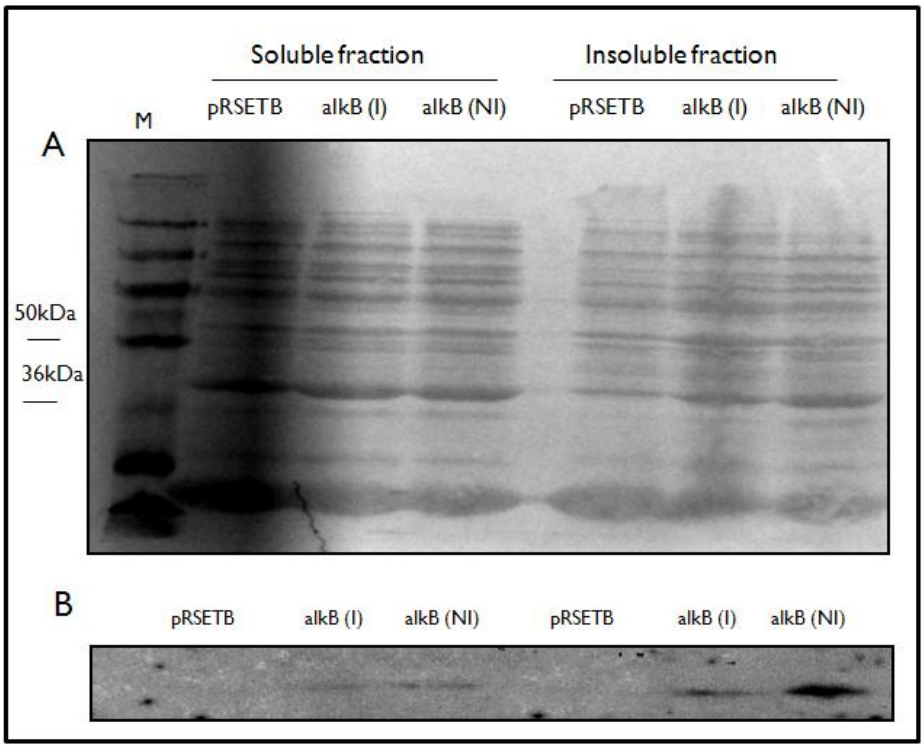

Figure 23. Expression of AlkB in *E.coli* BL21-AH. A) PAGE of total protein extracted after growth in Luria Bertani liquid medium with (I) and without (NI) IPTG as inducer from soluble and insoluble fraction. (pRSETB): protein extracted from the recombinat strain harbouring empty plasmid;(alkB): protein extracted from the recombinant strain harboring recombinant plasmid. B) Western blot analysis with anti-His tag antibody conjugated with alkaline phosphatase.

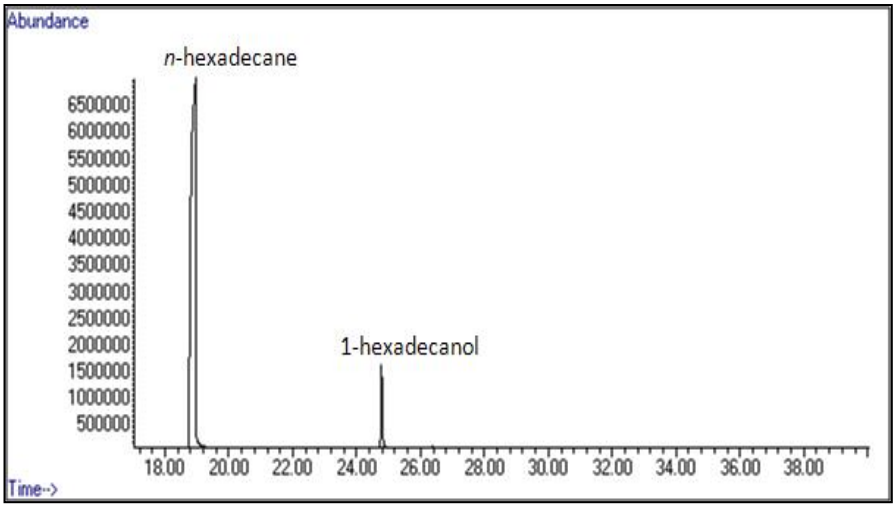

Figure 24. *n*-hexadecane bioconversion to 1- correspondent alcohol. GC-MS spectrum of solid phase microextraction (SPME) from cultures of *E.coli* BL21-AH after 22h of incubation in mineral medium supplemented with *n*-hexadecane.

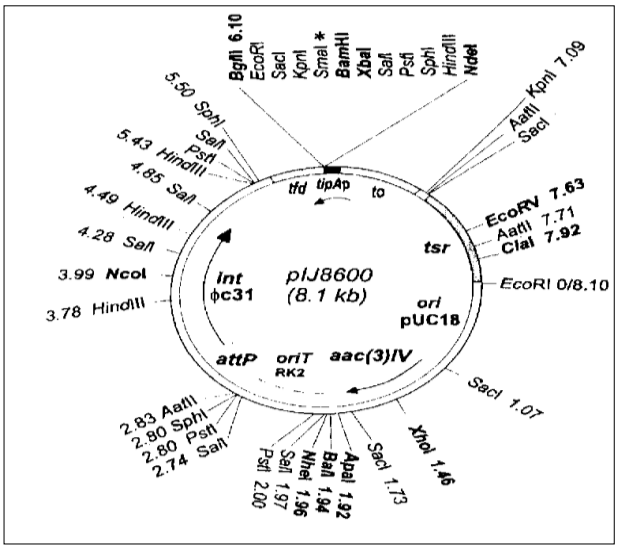

Figure 25. Restriction map of pIJ8600. *tfd*, major transcription terminator of phage fd; *to*, transcription terminator from phage λ; *tipA*p, the *tipA* promoter; *tsr*, Thio-resistance gene; *aac(3)IV*, apramycin-resistance gene selectable in *E. coli* and streptomycetes; *ori* pUC18, origin of replication from pUC18

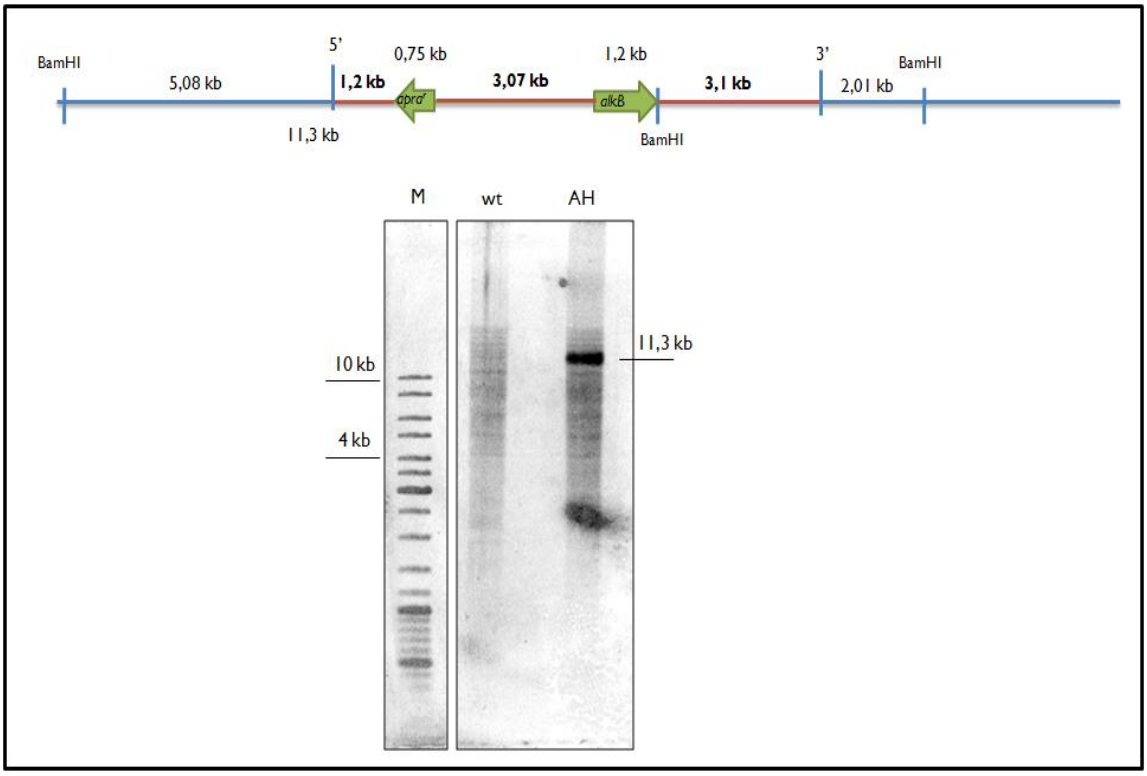

Figure 26. Schematic representation of the restriction map of genomic region after alkB integration and analysis of the plalkB correct *att-int* recombination into M145 genomic DNA. A) blu line: genomic region; red line: plalkB; apr^r : apramycin resistance cassette; 11,3kb is the size for 5'-3' recombination; 5,08kb is the size for 3'-5' recombination. Size of the more important genomic region were indicated. B)Overlay between agarose gel and film after chemiluminescent development. Southern analysis of the BamHI-digested genomic DNA of (wt) wild-type and (AH) recombinant strain. extracted from the (wt) wild-type *Gordonia* sp. strain SoCg and (Ω alkB) the mutant strain *Gordonia* SoCg, with DIG-labeled probe alkCg341. M: molecular weight marker.

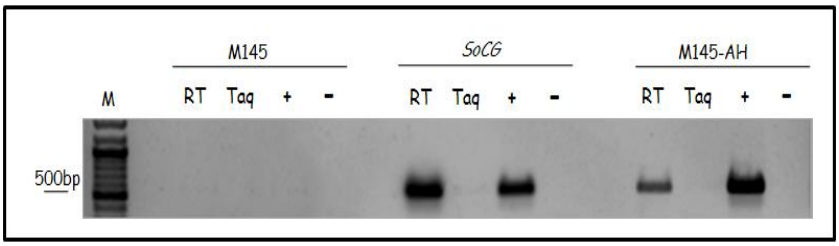

Figure 27. Heterologous expression of the *alkB* gene. Agarose gel electrophoresis of the amplification of *alkB* mRNA. Total RNA was extracted from (M145) *S.coeliclor* M145, (SoCg) *Gordonia sp.* strain SoCg, (M145-AH) *S.coelcolor* recombinant strain harboring integrative plalkB, after 48h incubation in rich medium with thiostrepton as inducer and retrotranscribed using RT-Taq (Invitrogen). (RT) amplification of RNA as template with the RT-Taq enzyme; (Taq) amplification of genomic DNA as template with the RT-Taq enzyme, this reaction was used to verify the absence of genomic DNA in the RNA sample; (+) positive control using genomic DNA of *Goronia sp.* strain SoCg as template; (-) negative control whitout DNA as template; M: molecular weight marker.

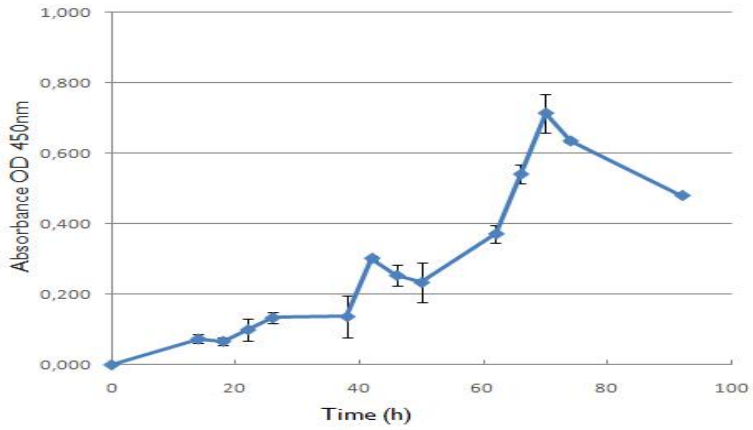

Figure 28. Time course of growth *of S.coeliolor* M145-AH on *n*-hexadecane. Standard Error vertical bars are calculated from three independent determinations.

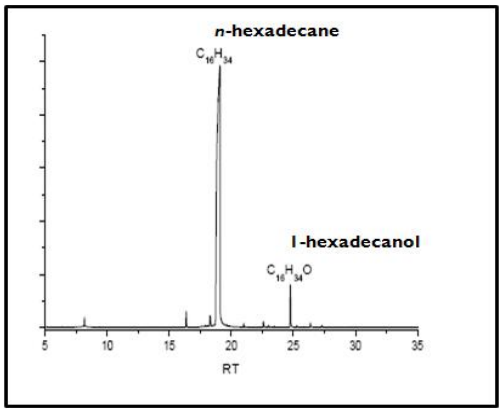

Figure 29. *n*-hexadecane bioconversion to 1- correspondent alcohol. GC-MS spectrum of solid phase microextraction (SPME) from cultures of *S.coelicolor* M145-AH after 22h of incubation in mineral medium supplemented with *n*-hexadecane.

7.4.5. *S. coelicolor* M145-AH expressing SoCg *alkB* grows on *n*-triacontane

In order to analyze the activity of *Gordonia* AlkB on solid *n*-alkanes we set a growth assay with *S. coelicolor* M145-AH in the presence of *n*-triacontane. *E. coli* was not used for growth assays because it lacks the metabolic pathway for the alcohol metabolization. The growth curve of M145-AH on triacontane (Fig.30) shows that gordonial *alkB* confers the ability to grow on solid *n*-alkanes and also that this strain possess in its own genome the genes involved in utilization of fatty alcohols. *S. coelicolor* M145 thus revealed to be a good system to study express long-chain *n*-alkane biotrasformation.

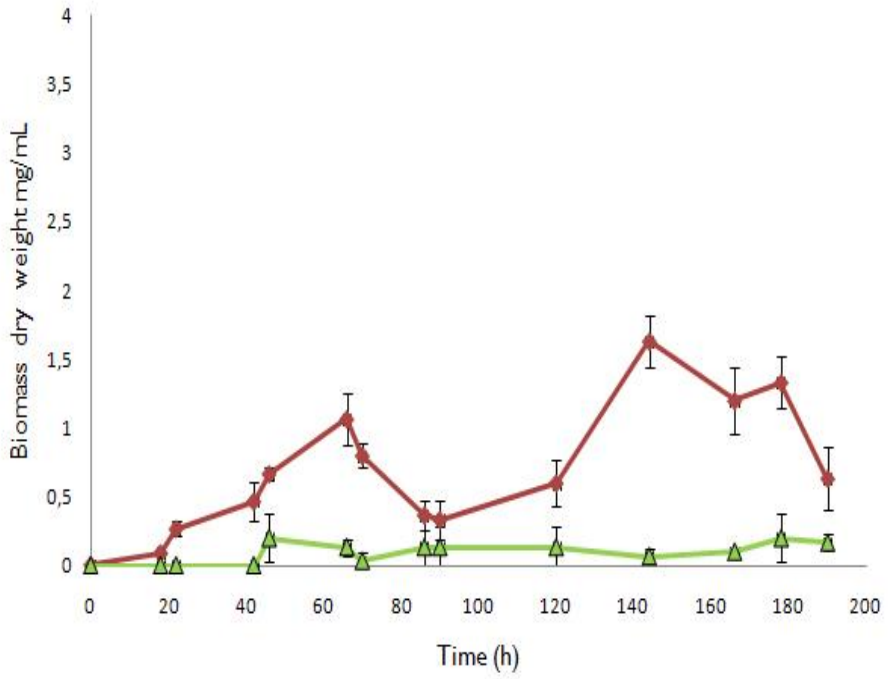

Figure 30. Comparison of biomass accumulation on *n*-hexadecane. Time course of growth of (●) *S.coelicolor* M145-AH and (▲) *S.coelicolor* M145. Standard Error vertical bars are calculated from three independent determinations.

8. Discussion

Many bacteria capable of degrading long-chain alkanes have been isolated, and the enzyme systems that oxidize long-chain *n*-alkanes up to C_{16} have been characterized (see Rojo 2009, van Beilen et al., 2007, Wentzel et al., 2007 for a review). Although long-chain alkanes are more persistent in the environment than shorter chain alkanes, genes involved in degradation of *n*-alkanes longer than C_{16} had not been reported prior to the works of Throne-Holst et al., (Throne-Holst et al., 2007) and Feng et al. (Feng et al., 2007). A flavin-binding monooxygenase involved in oxidation of very long chain *n*-alkanes up to C_{32} has been characterized in *Acinetobacter* sp. strain DSM17874 (Throne-Holst et al., 2007) and LadA from *Geobacillus thermodenitrificans* NG80-2 is the first long-chain *n*-alkane monooxygenase functional on alkanes in the range C_{15}–C_{36} to be cloned and structurally characterized from a Gram positive strain (Feng et al., 2007). Both enzymes show little or no homology with the widespread and well characterized AlkB-type alkane hydroxylases usually reported as functional on long chain *n*-alkanes up to C_{16} in Gram positive and Gram negative strains (Smtis et al., 2002; Wentzel et al., 2007; Whyte et al., 2002).

Here we describe for the first time the unique functional AlkB-type alkane hydroxylase system that allows growth on long chain liquid and solid *n*-alkanes in the Gram-positive *Gordonia* strain SoCg. Up to date the only long chain alkane hydroxylase system of this genus that has been characterized is that of *Gordonia* TF6 that was found to be functional on *n*-alkanes from C_5 to C_{13} (Fujii et al., 2004).

The ability of *Gordonia* SoCg AlkB to biotransform *n*-hexadecane into the corresponding primary alcohol was assessed by SPME/GC-MS analysis in SoCg and in two heterologous hosts expressing the SoCg *alkB*. *S. coelicolor* M145 was successfully used as heterologous host of an alkane-hydroxylase gene for the first time. Although we could not detect the *n*-triacontane biotransformation product, triacontanol, in none of the heterologous systems,

the role of *SoCg* AlkB in triacontane metabolization was demonstrated by growth assays: *S.coelicolor* M145-AH expressing the *Gordonia alkB* acquired the ability to grow on triacontane and the disruption mutant SoCg Ω*alk*B lost this ability. Moreover the SoCg alkane hydroxylase encoding genes are induced by both liquid and solid *n*-alkanes that is in accordance with the ability of this strain to grow on and rapidly metabolize *n*-alkanes up to C_{36} (Quatrini et al., 2008).

Taken together these results suggest that the identified alkane oxidation system plays a central role in the degradation of long and solid *n*-alkanes by *Gordonia* SoCg. Moreover it would exist at least another less efficient enzyme that is responsible for oxidation of *n*-hexadecane. This second AH system seems to be unrelated to other known alkane hydroxylase systems characterized so far.

Many bacterial strains contain multiple — quite divergent — integral membrane AlkBs (Whyte et al., 2007) that have different substrate ranges (Tani et al., 2001; Throne-Holst et al., 2007) or are activated during different growth phases (Marin et al., 2003). The strategy of *Gordonia* SoCg seems based on a single *alkB* gene, which is induced by a wide range of long and solid *n*-alkanes, throughout the entire time course of growth (L. Lo Piccolo, unpublished results), encoding an enzyme with highest activity on hexadecane and reduced activity on triacontane. For this reason growth of SoCg on triacontane would thus be poorer that that on hexadecane, and also because a second unknown system, functional on C_{16} but not on C_{30}, would contribute to overcome the limiting step of *n*-alkane degradation on C_{16}. The alkane-hydroxylase, in fact, catalyzes the initial attack and determines the size range of *n*-alkanes to be degraded; its specific activity is generally reduced with increasing chain length (Feng et al., 2007; Sameshima et al., 2008; Smits et al., 2002; Throne-Holst 2007; van Beilen et al., 2005).

The relationship between the AlkB protein structure and its function has been investigated; it has been proposed that AlkB is made of six transmembrane helices that are assembled in a hexagonal structure forming a deep hydrophobic pocket where four conserved histidine residues that chelate the iron atoms necessary for its activity, are located on the cytoplasm surface (van Beilen et al., 2005). The alkane molecule should slide into the pocket until the terminal methyl group is correctly positioned relative to the His residues. Amino acids with bulky side chain protruding into the pocket would limit the size of the *n*-alkane to be hydroxylated while less bulky side-chain amino acids allow longer alkanes to deeper enter into the hydrophobic pocket (van Beilen et al., 2005). *Pseudomonas putida* GPo1 and *Alkanivorax borkumensis* AP1 AlkB mutant derivatives oxidize alkanes longer than C_{12} when tryptophan is substituted by serine or cysteine or other small amino acids in position 55 or 58 of the two proteins (van Beilen et al., 2005). Amino acid sequence alignment of AlkB proteins showed a valine residue in the corresponding amino acid position of *Gordonia* SoCg AlkB, confirming the possibility to accept long-chain alkanes in the active site, although other residues/mechanisms could be involved in *n*-alkane recognition.

Bacteria appear to degrade chemicals only when they are dissolved in water and dissolution of solid substrates is generally considered prerequisite for their biodegradation (Wick et al., 2001). Long chain and solid *n*-alkanes are insoluble in water and, although we know how *n*-alkanes are oxidized, we still poorly know how they are recognized and how they enter the cells, especially when they are in solid state. Two mechanisms have been recognized in bacteria for accessing to medium and long chain liquid alkanes: 1) biosurfactant-mediated accession by cell contact with emulsified hydrocarbons and 2) interfacial accession by direct contact of the cell surface with the hydrocarbon (3). *Gordonia* belongs to the C*orynebacterium*/*Mycobacterium*/*Nocardia* (CMN) complex characterized by mycolic-acid containing cell walls that confer to these bacteria hydrophobicity and allow cell adherence to

the *n*-alkanes by direct contact of cells with hydrocarbons, generally with no or low biosurfactant production. Our observations confirm that the strategy of SoCg to access liquid hydrocarbons is by direct contact and that this strategy is also used for solid alkanes. In fact we observed a massive adhesion of SoCg cells to triacontane (supplemented as finely ground powder) while the culture liquid phase was almost clear for a long period of growth. Direct contact with the solid substrate might favour *Gordonia*'s growth as it can have direct access to the substrate without its previous solubilisation in the aqueous environment.

The first recent report of expression of *Rhodococcus alkB* genes in anhydrous organic solvents corroborates these observations (Sameshima et al., 2008) and suggests new biotechnological applications in water-free environments.

The alkane hydroxylase from *Gordonia* SoCg is not only active on a wide range of long chain liquid and solid *n*-alkanes, but revealed to be also a more versatile enzyme than *Pseudomonas*-type *n*-alkane 1-monooxygenases as it is able to use other electron transfer systems in the absence of its two specific components, rubredoxin and rubredoxin reductase. This versatility seems to be shared by other phylogenetically close AH from strains belonging to the *Corynebacterium-Mycobacterium-Nocardia* (CMN) complex (Fujii et al., 2004; Sameshima et al., 2008) strengthening the interest towards this group of bacteria.

Although hydrocarbon utilization is quite common among bacterial strains belonging to the CMN complex, only a few strains, mainly rhodococci, have been studied in some details. SoCg is the first *Gordonia* strain that is able to grow on solid *n*-alkanes to be characterized.

CHAPTER II

Converting excess carbon into storage materials

1. Microbial strategies for accessing long chain *n*-alkanes

Most substrates promoting microbial growth need to undergo cellular uptake or attachment to become accessible by the cell's catabolic machinery. This need, however, can be surpassed if hydrolytic enzymes (e.g., chitinases) are being secreted by bacteria and fungi to perform extracellular breakdown of substrates before uptake of the reaction products. In the case of water-insoluble substrates like *n*-alkanes, the hydrophobic nature of the bacterial cell surface has been reported to play an important role (Wentzel et al., 2007). The cell contact with hydrophobic substrates is crucial because the initial step in aliphatic and aromatic hydrocarbon degradation is often mediated by oxidation reactions catalyzed by cell-surface-associated oxygenases (Das et al., 2010). In the case of long-chain *n*-alkanes, two mechanisms for accessing these substrates are generally considered for bacteria: (1) interfacial accession by direct contact of the cell with the hydrocarbon and (2) biosurfactant-mediated accession by cell contact with emulsified hydrocarbons (Ron and Rosenberg, 2002; Van Hamme and Urban, 2009). Concerning surfactant-mediated accession, most of the *n*-alkane-degrading bacteria studied produce and secrete biosurfactants of diverse chemical nature that allow emulsification of hydrophobic compounds (Ron and Rosenberg, 2002).

It is generally believed that hydrocarbons interact with microorganisms nonspecifically and move passively into the cells. Of course, hydrocarbon-degrading microorganisms must necessarily come in contact with their substrates before any transport, either active or passive, may take place. Traditionally, three modes of hydrocarbon uptake are cited to describe how hydrocarbon-metabolizing organisms come in contact with their substrates. While microorganisms may contact water-solubilized hydrocarbons, decreasing solubility with increasing molecular weight is restrictive (Bury and Miller, 1993). Two additional, perhaps more widespread modes of hydrocarbon accession are direct adherence to large oil droplets and interaction with pseudosolubilized oil (Bouchez-Naïtali et al., 1999). For

example, Van Hamme and Ward (2001) described a *Rhodococcus* strain that grew directly on crude oil droplets and could be removed with the addition of exogenous chemical surfactant, while a *Pseudomonas* strain required surfactant-solubilized oil to efficiently access hydrocarbons. In *P.aeruginosa*, hydrocarbon solubilization and micellar transport control hexadecane biodegradation during biosurfactant-enhanced growth (Sekelsy and Shreve, 1999). Similarly, encapsulating solid *n*-C_{18} and *n*-C_{36} in liposomes increased growth and biodegradation by a *Pseudomonas* sp., indicating that cell-liposome fusion may deliver encapsulated hydrocarbons to membrane-bound enzymes.

With respect to active transport, proton motive force uncouplers have been shown to apparently decrease both *n*-hexadecane (Beal and Betts, 2000) and naphthalene (Whitman et al., 1998) uptake, which could indicate that energy-dependent uptake is important in some strains. In these two studies, the fact that the strains being studied could metabolize the substrates over the long incubation times complicates the separation of phenomena related to transport, metabolism, and growth. Probably the best observational evidence for energy-dependent *n*-alkane uptake is the case of *Rhodococcus erythropolis* S+14He, which preferentially accumulates *n*-hexadecane from hydrocarbon mixtures (Foght and Gray, 2002). Even though direct molecular evidence for active uptake has not been presented, it would not be surprising to find energy-dependent pumps that transport hydrocarbons into the cell. The presence of hydrocarbon inclusions, of both pure and partially oxidized alkanes, for example (Ishige et al., 2002), indicates that these substrates can be accumulated against a concentration gradient, presumably an energy-dependent process. In addition, as has been observed for 2,4-dichlorophenoxyacetate (Harwood et al., 1994) and 4-hydroxybenzoate (Hawkins and Harwood, 2002) metabolism, uptake and chemotaxis may be coordinately controlled at the molecular level.

1.1 Membrane Alterations

Given the hydrophobic nature of the area between the monolayers of the cytoplasmic membrane and, in gram-negative bacteria, of the outer membrane, it is not surprising that lipophilic molecules such as hydrocarbons partition there (van Hamme et al., 2003). Hydrocarbons tend to reside in the hydrophobic area between membrane monolayers in the acyl chains of phospholipids, with partitioning being related to the octanol-water partition coefficient of the lipophilic compound (de Bont and Poolman,1995). Hydrocarbon insertion alters membrane structure by changing fluidity and protein conformations and results in disruption of the barrier and energy transduction functions while affecting membrane-bound and embedded enzyme activity (de Bont, 1998).

In terms of general stress responses, bacteria may form biofilms, alter their cell surface hydrophobicity to regulate their partitioning with respect to hydrocarbon-water interfaces or, in gram-negative bacteria, gain protection from hydrophilic lipopolysaccharide components that offer high transfer resistance to lipophilic compounds. In addition, energy-dependent repair mechanisms may be used to compensate for losses in membrane integrity resulting from the partitioning of lipophilic compounds. For example, membrane fluidity can be decreased through increased membrane ordering by affecting *cis*/*trans* phospholipid isomerizations, by decreasing unsaturated fatty acid content, and by altering phospholipid head groups (Junkerand Ramos. 1999; RamosE. et al., 1997). These changes may be associated with an overall increase in phospholipid content and increased phospholipid biosynthesis in solvent-stressed cells (Pinkart, and White. 1997).

These alterations serve to produce a physical barrier to the intercalation of hydrocarbons in membranes, thus offsetting the passive influx of hydrocarbons into the cell.

2. Converting excess carbon into storage materials

When the carbon source is in excess relative to nitrogen, many bacteria transform part of the carbon into storage materials such as triacylglycerols (TAG), wax esters (WE), poly(hydroxybutyrate) or poly(3-hydroxyalkanoates) PHA, which accumulate as lipid bodies or as granules (Alvarez and Steinbüchel, 2002; Waltermann *et al*., 2005; Prieto, 2007). These neutral lipids can later serve as endogenous carbon and energy sources during starvation periods. Formation of storage lipids is frequent among hydrocarbon-utilizing marine bacteria. Alcanivorax strains, for example, can accumulate TAGs and WE when growing at the expense of pyruvate or *n*-alkanes (Kalscheuer *et al*., 2007). *Pseudomonas putida* GPo1, a soil bacterium, can form intracellular inclusions of poly-β-hydroxyoctanoate when grown on *n*-octane (de Smet *et al*., 1983), while *Acinetobacter* sp. M-1 forms wax esters from hexadecane (Ishige *et al*., 2000; 2002).

Thin sections of *Acinetobacter* sp. strain M-1 cells reveal electron-transparent intracellular inclusion bodies (Fig.1), which consist of wax esters. Quick-freezing replica microscopy more clearly shows the structure of the inclusion bodies, which are disk-shaped, have a smooth surface, and grow to almost the same diameter as the cells (Fig.1).

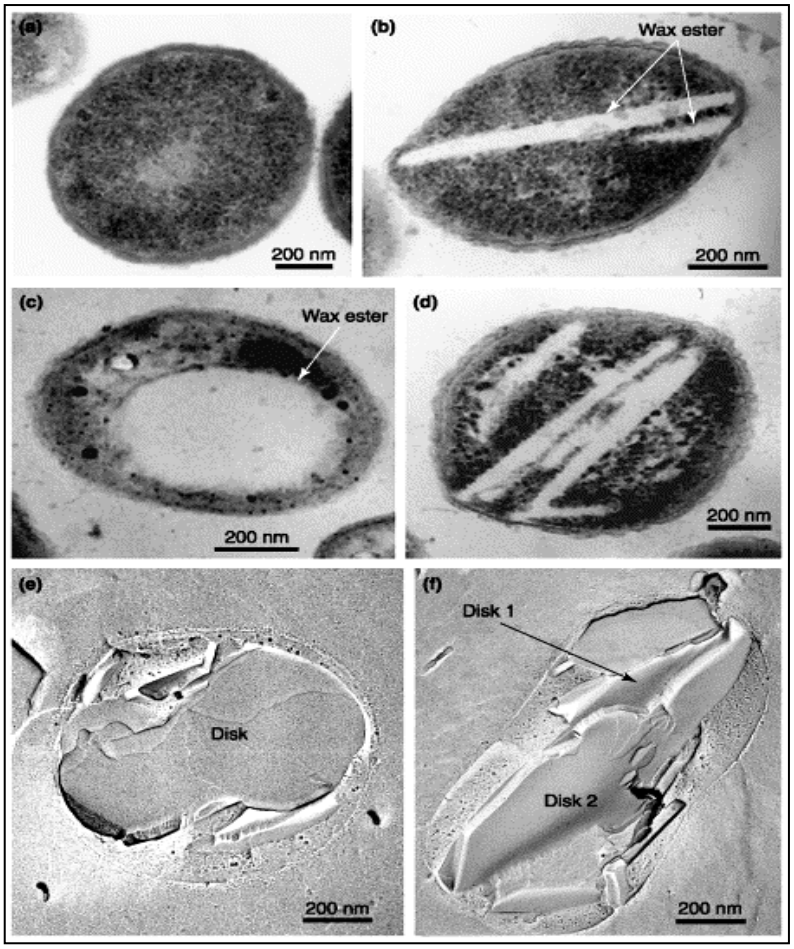

Figure 1. Electron microscopy of accumulated wax ester in *Acinetobacter* sp. strain M-1. Thin sections of cells were grown on (a) yeast extract-tryptone medium (2 × YT medium) and (b) after a shift to nitrogen-limited medium containing 0.5% *n*-hexadecane incubated for 2 h (b,c) and 10 h (d) at 30°C. Cells in (b and c) were incubated under the same conditions but sliced differently. (e,f) Images were obtained by the quick-freezing replica method of cells prepared under the same conditions as in (c). When cells grown on 2 x YT medium (a) were transferred into the medium containing 0.5% *n*-hexadecane (b), characteristic intracellular inclusion bodies containing wax esters were formed. The number of inclusion bodies increased during the incubation period (b,d). The inclusion bodies of wax esters were disk-shaped, had a smooth surface and grew to almost the same diameter as the cells (c,e,f). (from: Alvarez et al., 2008)

No intracytoplasmic membrane structures or limiting membranes surrounding these inclusions have been observed. The wax ester accumulation proceeds with the sequential formation of disks that are 30–50 nm in depth, (i.e. the completion of one disk leads to that of another).

In Acinetobacter M-1 WE were accumulated optimally up to 0.17 g/g of cells (dry weight) when cells (1.85×10^{10} cells/ml) were incubated with 3% *n*-hexadecane for 10 h at 30°C. The produced wax ester comprised hexadecyl hexadecanoate (98%) and hexadecyl myristate (2%). As shown in Fig 2. the maximum production of WEs occur during growth on medium chain length *n*-alkane up to C_{18}. Interestingly, the biosynthesis degree when the length of *n*-alkane increase, in fact only 0,09mg/mg cell of WE was produced during growth on C_{24}.

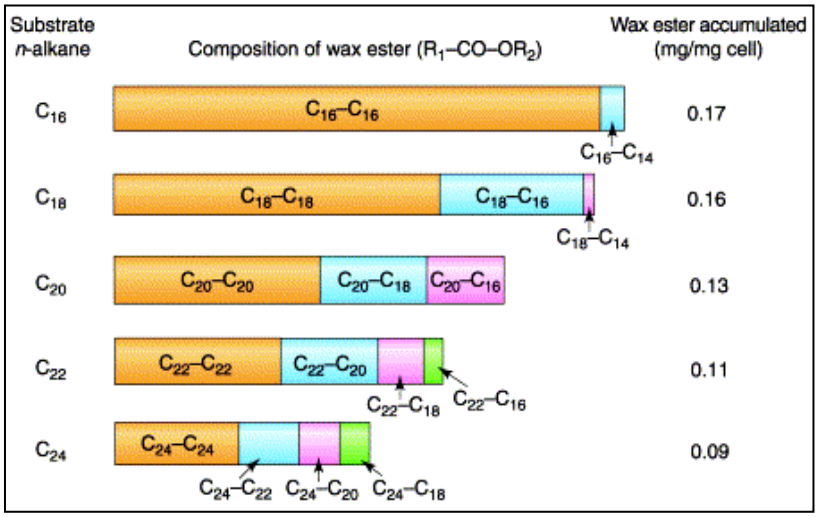

Figure 2. . Wax ester composition and production from various *n*-alkanes by *Acinetobacter* sp. strain M-1. The cells grown on yeast extract-tryptone medium were transferred to nitrogen-limiting media containing *n*-alkanes (C_{16}–C_{24}) and then incubated at 30°C for 10 h (from: Ishige et al., 2003)

2.1. Physiological role of the neutral lipids

TAGs are excellent reserve materials for several reasons. Their extreme hydrophobicity allows their accumulation in large amounts in cells without changing the osmolarity of the cytoplasm. Oxidation of TAGs produces the relatively high yields of energy in comparison with other storage compounds such as PHA and carbohydrates, since the carbon atoms of TAG acyl residues are in a very reduced state (Alvarez et al., 2002). Moreover, TAGs may play other important functions in cells of actinomycetes, including (1) regulating the fatty acid composition of lipid membranes, (2) as a sink for reducing equivalents in cells when the terminal acceptor is not sufficiently supplied under low oxygen conditions, (3) as source of precursors for biosynthesis and turnover of mycolic acids during adaptation to changing environmental conditions, (4) as a reservoir of metabolic water for cells under water stress conditions, (5) as an agent to detoxify free fatty acids or unusual fatty acids that may disturb membrane fluidity during catabolism of hydrocarbons, or (6) as a source of precursors for the biosynthesis of antibiotics (Banchio and Gramajo 2002; Silva et al., 2007).

2.2. Pathways for Wax ester synthesis

Condensation of acyl-CoA and diacylglycerol catalysed by a diacylglycerol-acyltransferase (DGAT or Atf) is the key enzymic step in WE biosynthesis, since reactions involved in the formation of the diacylglycerol substrate are also part of the phospholipid biosynthesis routes Fig.3.

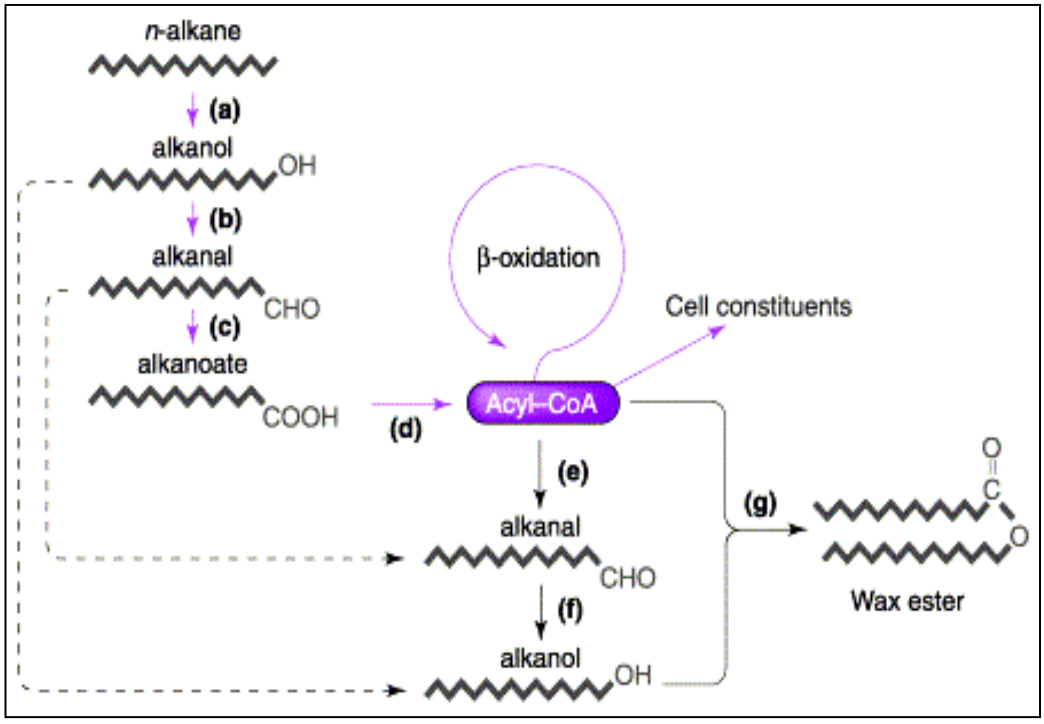

Figure 3. The proposed carbon flow from *n*-alkanes to wax esters in *Acinetobacter* spp (a) alkane monooxygenase, (b) alcohol dehydrogenase, (c) aldehyde dehydrogenase, (d) acyl–CoA synthetase, (e) acyl–CoA reductase, (f) aldehyde reductase (alcohol dehydrogenase) and (g) acyl–CoA:alcohol transferase.(from: Ishige et al., 2003)

The first Atf enzyme was described in *Acinetobacter baylyi* strain ADP1, and it exhibited WE synthase (WS) and DGAT activities (Kalscheuer & Steinbüchel, 2003). *A. baylyi* ADP1 possesses only a single Atf enzyme, which is responsible for synthesis of WEs as the main lipid storage compound in addition to minor amounts of TAGs (Kalscheuer & Steinbüchel,

2003). This enzyme exhibits a high promiscuity with respect to acyl acceptor molecules *in vitro* (Kalscheuer *et al.*, 2003; Stöveken *et al.*, 2005; Uthoff *et al.*, 2005). This is the reason why it is attracting increasing interest for biotechnological processes, for example 'MicroDiesel' production by recombinant *Escherichia coli* strains (Kalscheuer *et al.*, 2006).

2.3. Atf enzymes

Atf enzymes constitute a heterogeneous family of bacterial acyl-CoA-acyltranferases which do not have significant sequence similarities to eukaryotic DGAT or WS enzymes (Fig.4) (Wältermann *et al.*, 2006). Kalscheuer and Steinbüchel identified a bifunctional enzyme from *Acinetobacter baylyi sp.* ADP1 that exhibits simultaneously both acyl-CoA:diacylglycerol acyltransferase and acyl-CoA:fatty alcohol acyltransferase (wax ester synthase) activities (Kalscheuer and Steinbüchel, 2003).

Daniel et al. identified 15 putative WS/DGAT genes in *M. tuberculosis* strain H37Rv, which showed acyltransferase activity when expressed in *E. coli* (Daniel et al., 2004). Some of these genes, like rv3130c, show the highest induction and activity during hypoxia (Sirakova et al., 2006). Multiple *atf* genes were also identified in the genomes of other actinomycetes, including three in *S. coelicolor* (Arabolaza et al., 2008), 12 in *M. bovis* AF2122/97, 8 in *M. smegmatis* mc2155, 5 in *N. farcinica* IFM 10152 and 4 in *Nocardioides sp.* JS614 (Wältermann et al., 2007).

By contrast, TAG/wax esters accumulating gram-negative bacteria seem to have single *atf* genes, as is the case for *A. baylyi sp.* ADP1, *Psychrobacter sp.* and *Polaromonas sp.* (Wältermann et al., 2007). In general, actinomycetes accumulate higher amounts of TAG than gram-negative bacteria, and they have a high lipid content in different cellular structures, such as cell envelope or lipid inclusions (Alvarez, 2003).

A highly conserved motif HHxxxDG, which may be the catalytic site responsible for ester bond formation, is found in WS/DGATs from all known TAG-accumulating bacteria, but very low similarity exist in the whole gene sequence in respect to acyltrasferase sequence yet known.

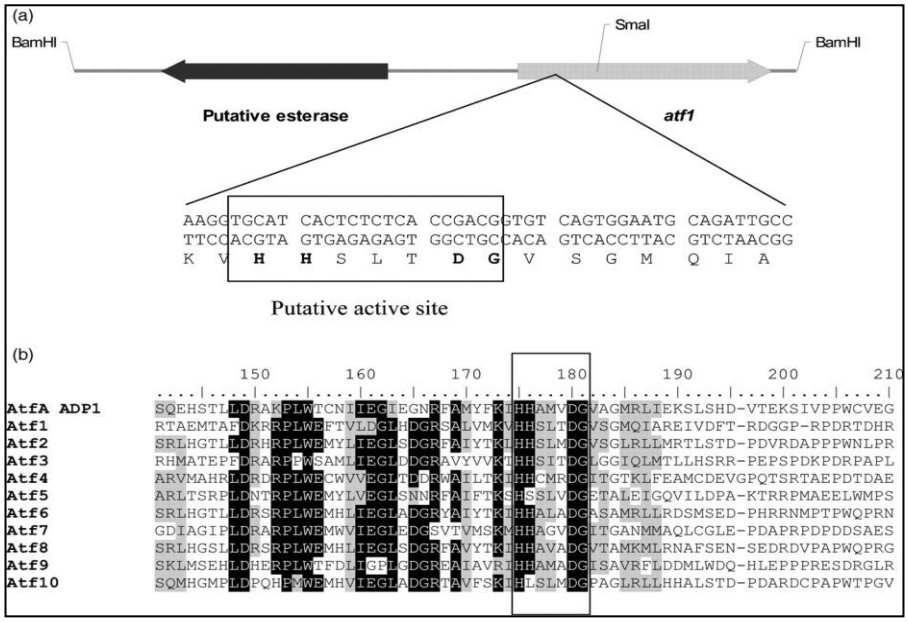

Figure 4. (a) Molecular organization of the 3948 bp *Bam*HI restriction fragment of *R. opacus* strain PD630 chromosomal DNA (fragment F11) harbouring the *atf1* gene. The sequence of the region of *atf1* that contains the putative active site is shown. (b) Multiple sequence alignment of the regions constituting the suggested active sites of Atf proteins from *R. opacus* PD630 with the AtfA protein from *A. baylyi* strain ADP1. Conserved amino acids are shaded in black, and homologous residues in grey. The putative active site is boxed.(from: Wältermann *et al.*, 2006)

Aims

The aims of this work were to investigate the presence of the acyltransferase gene in *Gordonia* SoCg by cloning and sequencing the *atfa* internal fragment; to analyze the *atfa*-like gene expression in relation to the *n*-alkanes consumption, and to explore the neutral lipids accumulation as products of *n*-alkane conversion.

3. Materials and Methods

3.1. Analysis of neutral lipid

3.1.1. TLC analysis

The neutral lipids as metabolic products of *n*-alkane bioconversion by *Gordonia* SoCg were analyzed by thin layer chromatography and solid-phase micro extraction (SPME) coupled with GC-MS. Neutral lipids of *Gordonia* SoCg was extracted from 2mg dried biomass after 48h growth on mineral medium with different *n*-alkane as carbon source, using 2mL chloroform/methanol (2:1, v/v) at room temperature for 3h. The solvents were evaporated from the combined extracts, and the residual material was dissolved in 50µL of the same solvent mixture. To determine the identity of the lipid and analyze the lipid profile from different growth conditions, TLC was carried out on silica gel 60 F_{254} plate (0,2mm, Mercks) loading an aliquot of 5µL of each sample obtained and using hexane/diethyl ether/acetic acid (80:20:1) as solvent. Lipid fractions were visualized by heating the plates over after staining in anisaldeide/sulfuric acid/ethanol and compared to Triolein and cetylpalmitate as TAG or wax ester reference substance, respectively.

3.1.2. SPME-GC-MS analysis

The entire suspensions were analyzed by immersing the SPME fiber coated with 85 µm polyacrylate (PA) and equipped with holder for manual injection. The time needed to reach equilibrium between the amount of analyte adsorbed by the polymeric film and the initial concentration of the analyte in the sample matrix, during the SPME sampling, is dependent on the properties of both the analyte and the matrix (de Pasquale et al., 2007) and in our study was 20 min at 45°C. Prior to use, the fiber was conditioned at 300 °C for 2 h, each in the GC injector port. HP-5MS 5% Phenyl Methyl Siloxane Capillary column was used to perform the gas-chromatographic separations. The oven initial temperature was 80 °C with

an helium constant flow, corresponding to the nominal head pressure of 9.37 psi. The temperature increased of 5 °C min^{-1} to 280 °C, temperature held for 20 minutes. The ionization spectra were obtained as mentioned above. The analytical identification and quantifications were carried out using standard grade compounds purchased from Sigma-Aldrich (USA) and commercial NIST 2005 mass spectra library search database.

3.2. Detection of *atfa*

The presence of *atfa-like* genes was detected using strategy PCR-based with different primers pairs. The first one degenerated primers pairs were designed on the basis of *atfa2* sequence of *Rhodococcus opacus* strain PD630 (*atfa2FR/RV*). Other approach was performed using degenerate primers pairs in the basis of consensus sequence obtained with alignment of the three putative acylglycerols enzymes of more phylogenetically related *Gordonia bronchialis* (Table 2).

Finally, the primers pairs as described by Alvarez et al., (2008), *tgs-int* (Table 2), were successfully used with this follows thermocycling parameters: 5 min at 94 °C, 30 cycles of 0.5 min at 94 °C, 0.5 min at 45 °C and 1 min at 72 °C, and finally 10 min at 72 °C.

The PCR product was cloned into pGEMT-Easy and subjected to DNA sequencing (Table 1).

Specific primer pairs *atfa* FR/ *atfa* RV (Table 2) were designed on the basis of sequence obtained from previously results and were used to specifically amplify an internal region of *atfa* gene, follow the previously described thermocyclung parameters.

Southern hybridization was carried out by using the protocols provided by Amersham. An *atfa* probe was obtained by labelled with the digoxygenin (DIG) system (Boehringer Manheim Biochemicals, Indianapolis, Ind.) the specific internal region of *atfa* and genomic DNA digested with with BamHI-EcoRI.

3.3. In vivo analysis of *atfa*-like gene expression

The cells were broken by using 3 mg of lysozyme/ml in P-buffer (Kieser et al., 2000) and total RNA was isolated follow the same procedures described in chapter 1. Reverse transcription-PCR (RT-PCR) was performed by using the Superscript One-Step RT-PCR kit (Invitrogen) with about 0.1 µg of total RNA as a template, primer pairs internal to *atfa*-like gene (*atfa* FR/ *atfa* RV) (Table 2) and the conditions indicated by the supplier, routinely using 35 PCR cycles. For each reaction, a negative control with *Taq* polymerase and without reverse transcriptase was included, in order to exclude DNA contamination.

4. Results

4.1. Analysis of storage compounds

Gordonia SoCg was grown for 48h in mineral medium supplemented by hexadecane, octadecane, eicosane, tetracosane, octacosane and triacontane, respectively, as sole carbon source and neutral lipids were extracted from the same amount of dry cells biomass and resuspended in the same volume of chloroform/methanol (2:1, v/v), in order to quantitatively compare storage compounds production and correlat their biosynthesis with *n*-alkane substrate. Neutral lipids extracted were spotted on thin layer chromatography (TLC) and analysis was conducted using hexane/diethyl ether/acetic acid (80:20:1) as carrier solvent and the same volume of (Fig.5).

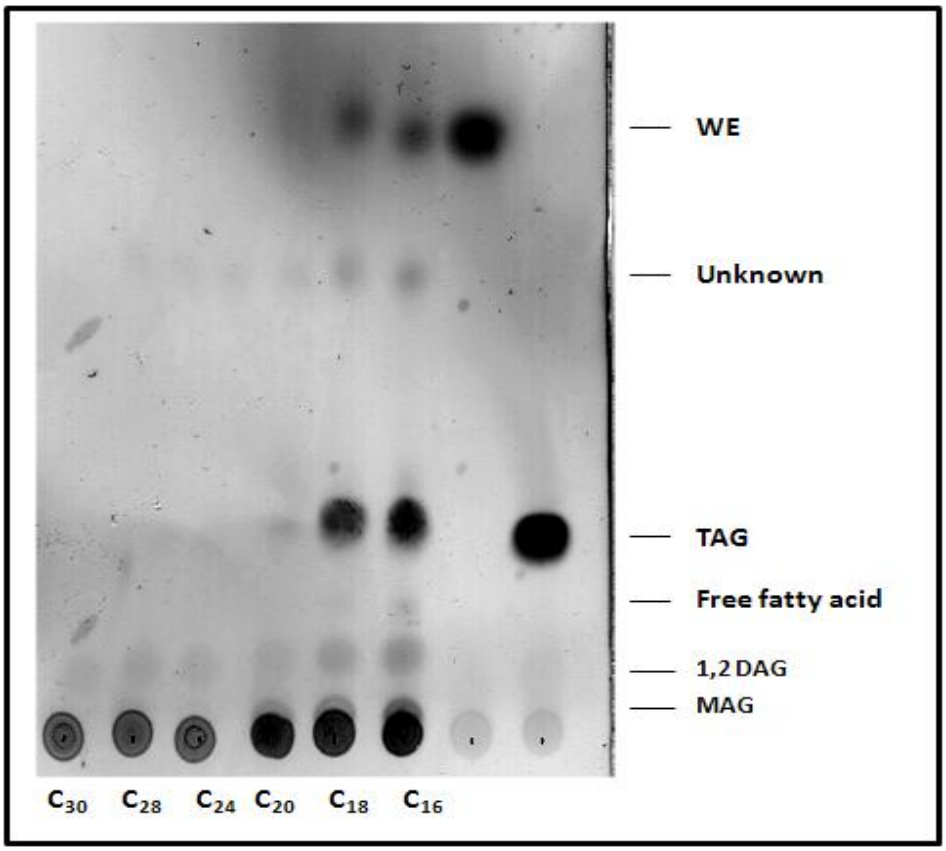

Figure 5. TLC Analysis of neutral lipids. Lipids were extracted from *Gordonia* SoCg after 48h of incubation in mineral medium supplemented with different *n*-alkanes as sole carbon source. Thin Layer chromatography was used to separate lipid fractions with *n*-hexane:dietyletere:acetic acid (80:20:1[vol/vol/vol]) as the solvent system. All molecules used as standard were 99% purities (Invitrogen). WE: cetylpalmitate as wax ester standard molecule; TAG: Triolein as standard molecule; 1,2 DAG: diacylglycerol and MAG: monoacylglycerol standards molecules.

This analysis indicated that *Gordonia SoCg* is able to produce different type of molecules as neutral lipids with different efficiency depending on the length of *n*-alkanes used as substrates, in according with those described by Ishige (2003). The most complete neutral lipids pattern was identified in sample obtained after growth in liquid *n*-alkanes, as hexadecane and octadecane and gradually decreases whit increasing length of the *n*-alkane. TAGs and WEs were identified as the most representative neutral lipids products, suggesting an interesting ability for the production of oleaginous compounds in this strain.

In order to explore if the absence of neutral lipids in the sample obtained from solid alkane growth condition was depending on low threshold of detection of the methods used, SPME analysis coupled with GC-MS was performed from neutral lipids collected from triacontane growth (Fig.6).

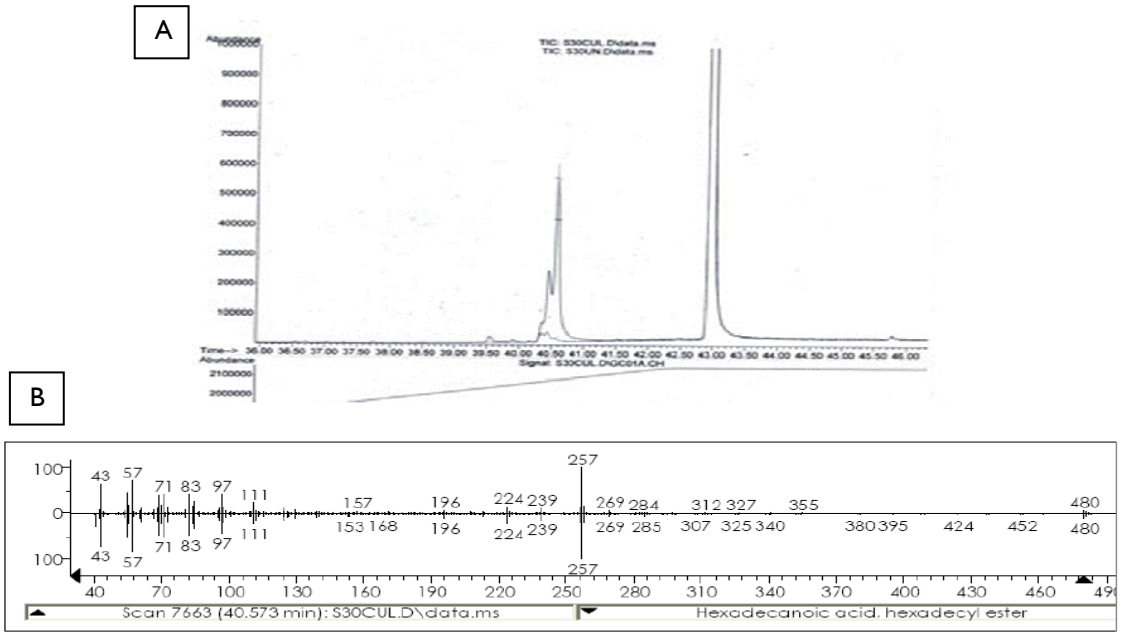

Figure 6. Chromatogram of wax esters of *Gordonia SoCg* grown on *n*-triacontane. Analysis was carried out by GCMS analytical technique by using a Hewlett-Packard 5890 GC system interfaced with an HP 5973 quadrupole mass spectrometer detector. As stationary phase an HP5-MS capillary column (5% 11 diphenyl – 95% dimethylpolysiloxane 30 m – 0.2 mm, 0.25 μm film thick, J&W Scientific, USA) was used. (A) Total ion chromatogram showing WE product peak at a retention time of 41.9 min; (B) Electron impact mass spectrum of WE product peak. The masses of the parent ion at 590 *m/z* and fragment ions (365, 311, and 278 *m/z*) match those reported for the hexadecyl hexadecanoic acid.

Solid phase microextraction was carried out using biomass of *Gordonia SoCg* after 48h of incubation in the presence of *n*-triacontane as sole carbon source. Interestingly, this most sensitive and accurate analytical technique revealed a peak corresponding to tetracosanoic acid butyl ester (RT 40.576), that was absent in the abiotic cultures (Fig.6), suggesting that wax ester were produced also during growth on solid *n*-alkane, but in lower amount.

4.2. Identification of *atfa*-like gene in *Gordonia SoCg*

Given that *Gordonia sp*. strain SoCg is able to convert *n*-alkane in neutral lipids, this strain was screened for genes coding for this metabolic activity using a PCR-based strategy and different primer pairs. It is known that the enzymes acyltransferases (AtfA) in bacteria are responsible for esterification of primary long chain to TAGs and WEs. Different primers pairs were used to identify an acyltransferase gene sequence in the genome of *Gordonia sp*. strain SoCg (Table 2). On the basis of acyltransferase 2 gene sequence of the best characterized *R. opacus* PD630 one primer pair was designed (*atfa2FR/RV*) and used in a PCR reaction without significant results. Another approach was conducted that was similar to that successfully applied by Alvarez (Alvarez et al., 2008) in *atfa* gene identification in *R. opacus* PD630, that required sequence analysis of the most phylogenetically related strain whose genome sequence is known. To date, within

the *Gordonia* genus, only the *Gordonia bronchialis* DSM 43247 genome is completely sequenced (GenBank: CP001802.1). Using the 14 *atfa* gene sequence of *R. jostii* RHA1 the *Gordonia bronchialis atfa*-like gene were explored using BioEdit local alignment (BioEdit v.7.0.5 http://www.mbio.ncsu.edu/bioedit/bioedit.html) and at least three genome regions were identified.

The diacylglycerol O-acyltransferase gene, the more strictly related acyltransferase 1(at genome position 262204163), and the acyltransferase 11 (at genome position 262203771) were aligned by ClustalX (http://www.clustal.org/) (Fig.3). Interestingly, all the three sequences contain the HHxxxDG putative site at positions 135, 139,138, respectively of the amino acid sequence (Fig. 3).

```
                 ....|....| ....|....| ....|....| ....|....| ....|....|
                         110        120        130        140        150
4163             LRYVSLNHGA LLDRFRPMWE FHIIEGLDDG RVALYSKLHH SLVDGVSALR
3771             AELVGHLAGQ TLDRGKPLWE LWVIEGSSDG RITAMLRMHH AGTDGVTSAE
diacyl           DAVLADLAVR QLDRSRPLWA LTLVHGLAGG RQAVVVRVHH AVADGLAALN
Clustal Co          :.       *** :*:*  : ::.*  .* * :   ::** : .**::: .
```

Figure 7. Aminoacid alligment of the three atfa-like gene from Gordonia bronchialis DSM 43247

Thus, three different primers pairs were designed (diacylFR/RV, 4163FR/RV, 3771FR/RV) (Table 2) and used in PCR reaction, but unfortunately, without results.

Finally, A 1234 bp PCR product was obtained from chromosomal DNA of strain SoCg by using degenerate primers *tgs-int 1/2* designed by Alvarez on the best characterized atfa gene of *Rhodococcus opacus* PD630 (Fig.4) (Table 2). This amplicon showed a molecular size about 400 bp larger than the expected one. Sequence of the amplicon (cloned in TOPO-TA vector; see materials and method- Table 1) showed 26% similarity to *wx/dgat* from *Acinetobacter sp.* ADP1 and it doesn't show the **HHxxxDG** putative catalytic site.

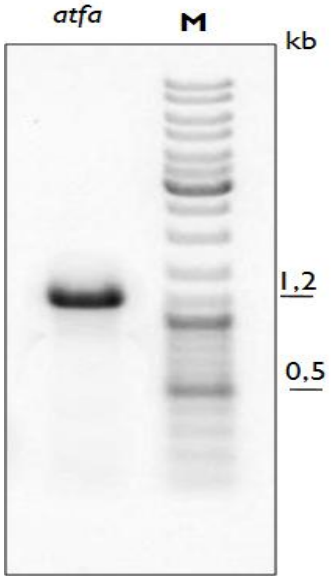

Figure 8. Agarose gel electrophoresis analysis of *atfa* amplicone from *Gordonia SoCg*, obtained using Alvarez's primers pairs.

4.3. Screening for estimation of copy number of the *atfa*-like gene

Specific primer pairs were designed on an internal region of the 1234 amplicone obtained with Alvarez's primers pairs (Table 2) to obtain a 774bp fragment to be used as a probe in Southern hybridization experiments in order to verify how many copies of this gene are present in *Gordonia SoCg* chromosome. The SoCg DNA was digested with several single and coupled enzymes and probed with the *atfa*774bp fragment. Surprisingly, only one fragment was revealed in all cases (Fig.5). This is a quite unexpected result because other phylogenetically related bacteria usually contain at least three homologs of this gene (Wältermann *et al.*, 2006).

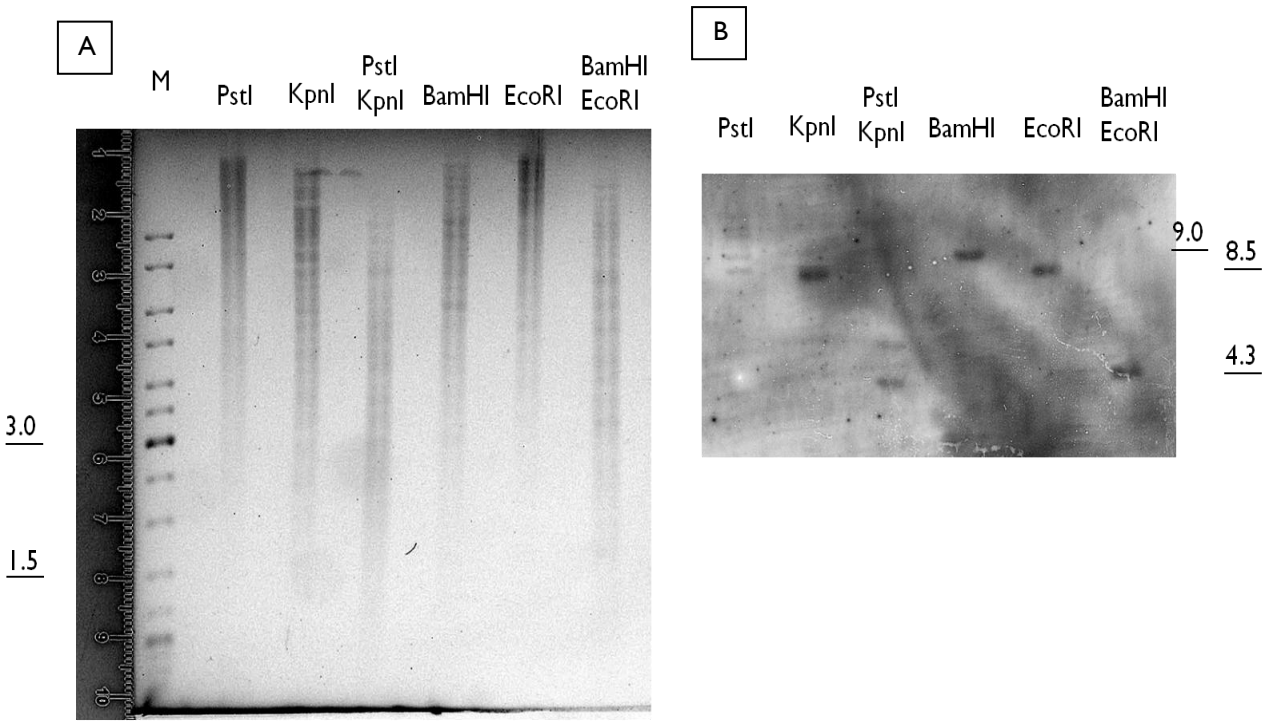

Figure 9. Southern analysis of the genomic DNA digested with different restriction enzymes. A: agarose gel electrophoresis of genomic DNA digested wit PstI, KpnI and coupled PstI-KpnI enzymes and of genomic DNA digested with BamHI, EcoRI and coupled BamHI-EcoRI enzymes. B: southern analysis with DIG-labeled probe atfa774.

In order to isolate entire *atfa* gene, an enriched gene bank was constructed in *E.coli* DH10B using by isolating 4,5Kb BamHI-EcoRI restriction fragments (Fig.9) from a preparative gel. Genome library of 476 recombinants clone is under analysis.

4.4. In vivo expression analysis of atfa-like gene

To study the influence of long chain *n*-alkanes on the expression of *atfa*-like gene, *Gordonia SoCg* was cultivated in a mineral salt medium in the presence of *n*-hexadecane and *n*-triacontane, as sole carbon source and total RNA was isolated after 22h of incubation at 30°C. RT-PCR analysis revealed that *atfa*-like gene is induced by both liquid and solid *n*-alkanes and it is not transcribed in the presence of fructose as sole carbon source (Fig.10). This result suggest that this gene probably plays an important role in wax ester synthesis.

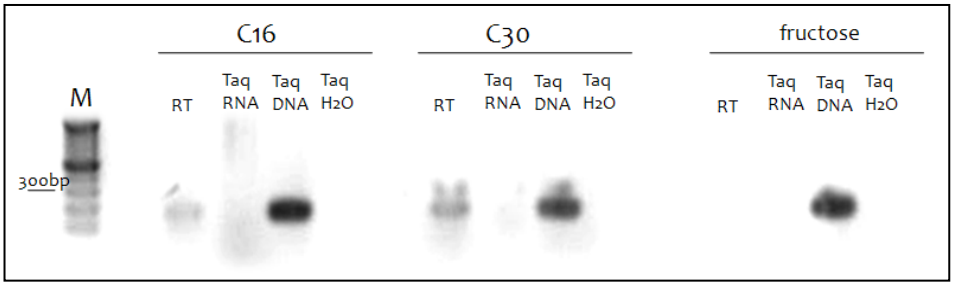

Figure 10. Agarose gel electrophoresis of the amplification of *atfa*-like mRNA. Total RNA was extracted from *Gordonia sp.* strain SoCg after 22h incubation in mineral medium supplemented with C16: *n*-hexadecane; C30: *n*-triacontane; fructose: fructose, as the sole carbon source, respectively. and retrotranscribed using RT-Taq (Invitrogen). (RT) amplification of RNA as template with the RT-Taq enzyme; (Taq RNA) amplification of RNA with the RT-Taq enzyme, this reaction was used to verify the absence of genomic DNA in the RNA sample; (Taq DNA) positive control using genomic DNA of *Goronia sp.* strain SoCg as template; (TAq H_2O) negative control without DNA as template; M: molecular weight marker.

5. Discussions

It is known that when the carbon source is in excess relative to nitrogen, many bacteria transform part of this into storage materials such as triacylglycerols, wax esters, poly(hydroxybutyrate) or poly(3-hydroxyalkanoates), which accumulate as lipid bodies or as granules.

The recalcitrant liquid and solid *n*-alkane are oxidize to long chain fatty alcohol, as previously described in chapter 1 and finally, completely metabolize via oxidative process to fatty acid needed to support growth of the cell. It was demonstrated (Ishige et al., 2003) that part of the 1-alkanol is esterified with fatty acid CoA thioesters by the enzyme WS/DGAT (Wältermann et al., 2007). This reaction generate high energy value molecules, that are packed in internal vesicles as storage compounds. This phenomenon was descrive in *Acinetobacter* and in *Alcanivorax* strains and also in the oleaginous bacteria such as *Rhodococcus*. In this work, it was found that also *Gordonia sp.* strain SoCg is able to produce neutral lipids during growth in the presence of *n*-alkanes as the sole carbone source and in particolar, by TLC analysis, TAGs and Wes were identified as major products of them.

The accumulation of storage lipids by actinomycetes, like members of *Mycobacterium*, *Rhodococcus*, *Nocardia* and *Streptomyces* is a well-established feature (Alvarez 2006).

The first enzyme involved in WEs and TAGs biosynthesis was described in *Acinetobacter baylyi* strain ADP1, and it exhibited condensation of acyl-CoA and diacylglycerol activity (Kalscheuer & Steinbüchel, 2003). A highly conserved motif HHxxxDG, which may be the catalytic site responsible for ester bond formation, is found in WS/DGATs from all known TAG-accumulating bacteria, but very low similarity exist in the whole gene sequence respect to acyltrasferase sequence yet known.

The multiplicity of *atf* genes in *Rhodococcus* and other actinomycetes, and their ability to synthesize TAGs as main storage lipids, suggest an important role of these lipids in the physiology of these organisms and their ability to cope with adverse environments.

Using degenerated primers pairs, designed by Alvarez (Alvarez et al., 2008) on the basis of acyltransferase gene sequence of phylogenetically related bacteria, amplicone with unexpected molecular size was amplified and sequenced. The highly conserved motif HH*XXX*DG corresponding to amino acids 132–138 of the *A. calcoaceticus* ADP1 WS/DGAT, putative catalytic site responsible for ester bond formation, was not identified in this sequence. All primers pairs designed used in this work (see results; Table 2) were unfortunately failure.

The unique amplicone obtained show an unespected molecular size, 400 bp larger than expected, and very low nucleotide similarity with other *ws/dgat*. In any case it doesn't contain the putative catalytic site.

Further unexpected result was obtained in Southern blot analysis, with which only one copy of this putative acyltransferase was revealed in genome. By contrast, TAG/wax esters accumulating gram-positive bacteria seem to have more of one and at least three *atf* genes, and only gram negative-bacteria have only one copy, as previously reported.

These evidence could suggest a different role for this gene in SoCg strain or that the identified genome sequence could be a new acyltransferase gene, the first isolated in this poorly explored genus.

Nevertheless, the expression of putative *atfa* during growth on *n*-alkane, as sole carbone source, does lean toward the second hypothesis. In fact, both liquid and solid *n*-alkanes induce the expression of *atfa*, whereas it is transcribed at low levels when *Gordonia* grows on simple sugars, indicating a specific role in the alchol esterification, during growth on *n*-alkanes

In *Gordonia sp.* strain SoCg, the *n*-alkanes are bioconverted both to generate energy for cell proliferation and in storage compounds for cell survival. It was found that different *n*-alkane are biotransformed with different efficiency, because liquid *n*-alkane support rapid growth and high production of WEs and TAGs, instead low biomass is accumulated with solid *n*-alkane and very low amount of neutral lipids are produced. It is possible that different cell response to *n*-alkane could depend on oxidative AlkB activity on these substrate. As mentioned above, AlkB could better oxidize liquid than solid *n*-alkanes. According to this hypothesis, the large amount of long alcohol could be used to bioconversion also in storage compound, instead the very long alcohol produced on C_{30} could be exclusively used for growth sustainability. Other experiments are needed to verify these hypothesis, for example it would be useful to study the enzymatic activity of the alkane monoxygenase with different *n*-alkane as substrates and also it would be useful to analyze the enzymatic activity of acyltransferase with different alcohol as substrates. These analysis are needed to understand the metabolic destiny of *n*-alkane in bacteria and the physiological role of bacteria fattening in adverse environmental conditions.

CHAPTER III

Proteomic insights into metabolic adaptation to n-alkanes

1. Behavioral and physiological responses to hydrocarbons

The molecular and biochemical basis of microbial physiological responses to hydrocarbons and the impact of these responses on bioremediation have been neglected until very recently (van Hamme et al., 2003). In fact, the metabolic pathways driving the activation of hydrocarbons into central metabolic pathways are well understood, while behaviors and responses are not appreciated at a general observational level. However, these phenomena are essential for allowing hydrocarbon-metabolizing organisms to avoid toxic effects, to access poorly soluble substrates, and, in some cases, to bring very large substrates into the cell.

Only recently, at the global levels, physiological adaptation of bacteria in the presence of aliphatic hydrocarbon were studied, and in particular only three papers have been published in which were used proteomic approach to identify proteins involved in *n*-alkane metabolism in *Alcanivorax borkumensis*, M*arinobacter hydrocarbonoclasticus* SP17 and *Geobacillus thermodenitrificans* NG80-2 (J.S. Sabirova et al., 2006; P.J. Vaysse et al., 2009 ;L.Feng et al., 2007). Oil-degrading marine bacteria have been isolated from different sites all over the world (Head et al., 2006; Yakimov, 2007).

The low solubility of alkanes in water led hydrocarbonoclastic bacteria to develop effective uptake systems. As previously described, interfacial contact and emulsification or pseudo-solubilization are two strategies developed to overcome the low mass transfer rate of long chain alkanes to the bulk aqueous phase. In some case biofilm formation at the hydrocarbon–water interface might be an alternative strategy for overcoming the low water solubility of hydrocarbons (Johnsen A.R. and Karlson U. 2004).

Efficient transporter systems for uptake of nutrient elements and detoxification play a crucial role for the survival of bacteria under reservoir conditions and proteins involved in nutrient transport across cellular envelopes constitute a large class of proteins overproduced. This

includes components of phosphate and thiosulfate, exclusively in *M. hydrocarbonoclasticus* SP17, and members of the ATP-binding cassette transporter superfamily, and electrochemical potential-driven transporters for *A.borkumensis* and *G.thermodenitrificans* NG80-2.

Cytoplasmic proteins up-regulated in *n*-hexadecane-grown cells reflect a central metabolism based on a fatty acid diet, such as enzymes of tricarboxylic acid cycle TCA cycle and glyoxylate bypass and of the gluconeogenesis pathway, able to provide key metabolic intermediates, like phosphoenolpyruvate, from fatty acids. They also include enzymes for ancillary functions included the lipoprotein releasing system (Lol), presumably associated with biosurfactant release, and polyhydroxyalkanoate synthesis enzymes associated with carbon storage under conditions of carbon surfeit.

Most of the enzymes corresponding to the CO_2-releasing steps of the TCA cycle were found to be downregulated in *M. hydrocarbonoclasticus* SP17 during growth on hexadecane and without significant change in expression profile in the case of *A.borkumensis* and *G.thermodenitrificans* NG80-2. Interestingly, the gene encoding malate synthase catalyzing the conversion of glyoxylate into malate, was found upregulated in all three bacteria studied. This modulation of TCA cycle enzymes suggests stimulation of the glyoxylate bypass in BH (Fig.1).

The explanation of glyoxylate bypass stimulation could lie in the fact that hexadecane is a more energetic substrate than acetate. In fact, the breakdown of hexadecane into acetyl-CoA produces energy, whereas conversion of acetate to acetyl-CoA does not. During growth on hexadecane an increased flux of carbon through the glyoxylate bypass could enable cells to restore the balance between carbon assimilation and energy production.

Fatty acid biosynthesis genes were found to be downregulated in *M. hydrocarbonoclasticus* SP17 . This corroborates the fact that, in cells growing on hexadecane, the main fatty acids of cellular lipids are derived from oxidation of alkanes.

Proteome pathways involved in *n*-hexadecane metabolism in *G. thermodenitrificans* NG80-2 was proposed by Feng and colleagues (Feng et al., 2009) (Fig.2).

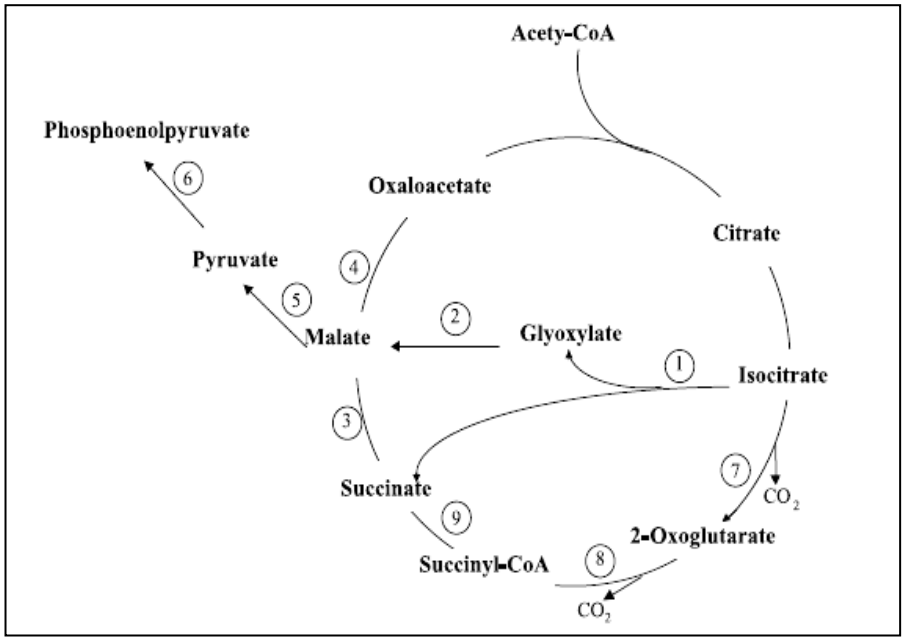

Figure 1. Increased exploitation of the glyoxylate bypass of the TCA cycle in alkane-grown cells of *Alcanivorax*. The glyoxylate bypass is carried out by isocitrate lyase (1) and malate synthase (2). Succinate produced via glyoxylate bypass is converted to malate by succinate dehydrogenase (3). Malate is converted to oxaloacetate by malate dehydrogenase (4) or is used by malic enzyme (5) in gluconeogenesis to produce pyruvate. Pyruvate is then converted by phosphoenolpyruvate synthase (6) to produce phosphoenolpyruvate. The incomplete TCA cycle is associated with the alkane-induced down-regulation of isocitrate dehydrogenase (7), 2-oxoglutarate dehydrogenase (8), and succinyl-CoA synthetase (9).

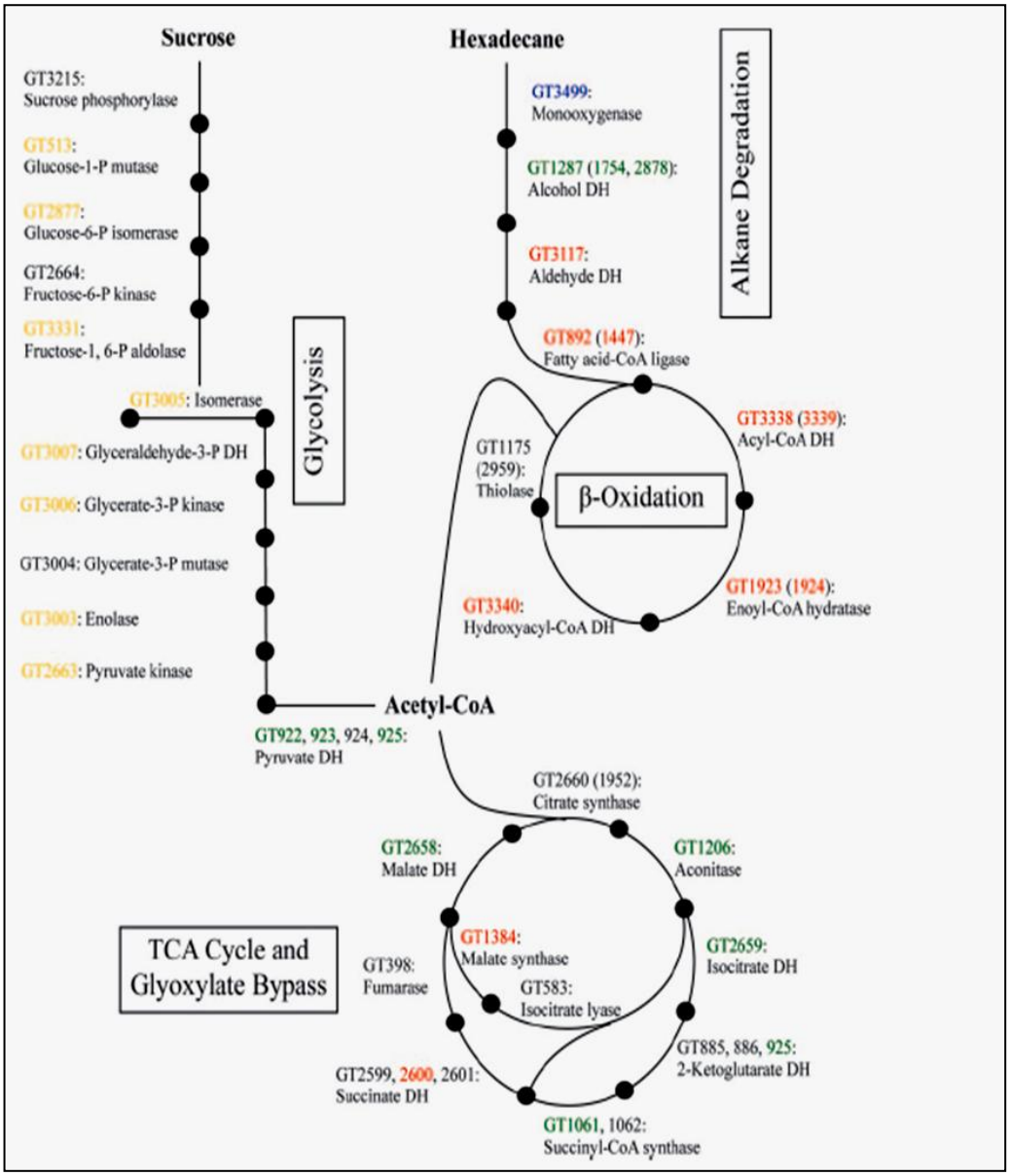

Figure 2. Proteomic characterization of pathways involved in hexadecane metabolism in *G. thermodenitrificans* NG80-2. Differentially expressed proteins in hexadecane-grown cells, in comparison with sucrose-grown cells, were investigated by 2-D electrophoresis (2-DE)/MALDI-TOF MS analysis. The predicted metabolic pathways for sucrose and hexadecane are shown. The enzymes of the pathways and corresponding genes in NG80-2 (as gene ID numbers) are noted. Proteins detected by 2-DE/MALDI-TOF MS are highlighted. Blue, induced; red up-regulated; yellow, down-regulated; green, unchanged. Proteins with pI values outside the range of pH 4 to 7 were not investigated. Analogs are in brackets. P, phosphate; DH, dehydrogenase.

2. The genus *Streptomyces*

The genus *Streptomyces* consists of Gram-positive, aerobic bacteria abundant in most soils, characterized by a complex life cycle, involving hyphae differentiation and spore formation. These microorganisms exhibit a mycelia growth habit and rely for their nutrition on a variety of extracellular hydrolytic enzymes such as cellulases, hemicellulases, proteases and lipases. These proteins allow *Streptomyces spp.* to use a wide range of complex insoluble substrates such as cellulose, hemicelluloses and oils as carbon source.

Their versatility allows them to live and propagate under unfavourable circumstances, such as in oil-polluted soils. It is therefore surprising that their use for bioremediation of hydrocarbon-polluted areas has been neglected. There are few publications about their use in hydrocarbon transformations, most of them restricted to the application in screening experiments (Sorkhoh et al. 1995).

In a short review (Barabas et al., 2001) it is discussed the recent research on oil utilization by *Streptomyces* strains (Barabás et al. 1995; Radwan et al. 1998) isolated from Kuwait oil fields (KCC strains). Most of the strains were identified as *S. griseoflavus*, *S. parvus*, or *S. plicatus* (Barabás et al. 2000). The strains were incubated with *n*-alkanes and increase of the fatty acid content with chain length equivalent to the employed *n*-alkanes was observed. Signal transducing GTP-binding proteins (GBPs) play an important role in *n*-alkane uptake in streptomycetes. Specific activators of GBPs increased the uptake of hydrocarbons. Using the hydrophobic fluorescent dye diphenylhexatrien (DPH) as a probe, it was found that the microviscosity of the hydrophobic inner region of the cellular membrane is significantly lower in hydrocarbon utilisers than in non-utilisers. This difference probably reflects differences in the fatty acid composition of the strains. When cultures were grown in *n*-alkane containing media, electron microscopy revealed that the hydrocarbon utilisers showed less-electron dense areas as inclusions in the cytoplasm. Soil samples inoculated with

Streptomyces strains eliminated hydrocarbons much faster than those not containing these strains, serving as control. When inorganic medium was supplied with *n*-hexadecane-1-^{14}C as sole carbon and energy source, radioactive CO_2 was detected.

Streptomycetes are also well able to utilize more readily available materials, such as simple sugars. Glucose is often preferred to other carbon sources, and as was shown in *S. coelicolor* A3(2), this preference is exerted either trough repression of genes involved in the uptake of alternative carbon sources, including arabinose, glycerol and galactose, or by repressing the production of extracellular polysaccharide-degrading enzymes such as agarase, α-amilase and chitinase.

S. coelicolor strain M145 is able to utilize exogenous fatty acid of different chain length, from C_4 to C_{18} as sole carbon and energy sources. It was found by Banchio C. and Gramajo H.C., that enzymes of the β-oxidation cycle are not inducted by substrates, but all five of them (acyl-CoA synthase, acyl-CoA dehydrogenase, enoyl-CoA hydratase, 3-hydroxyacyl –CoA dehydrogenase and 3-ketoacyl-CoA thiolase) involved in this catabolism were constitutively expressed, (Banchio C. and Gramajo H.C. 1997). Interestingly, it was also found by the same authors, that β-oxidation pathway in *S. coelicolor* strain M145, instead of being repressed by glucose, as known for genes involved in the uptake and/or utilization of many alternative carbon source, was stimulated by this metabolite.

Aims

The aims of this work were to understand metabolic adaptation to *n*-alkane polluted environments and to identify proteins involved *n*-alkane utilization.

3. Materials and Methods

3.1. Heterologous expression of *alkB* in *S. coelicolor*

The *alkB* gene was amplified using primers alkBamHIfor and alkBNdeIrev (primers table) from *Gordonia SoCg* genomic DNA and cloned in pGEM-T-easy vector (Promega). The resulting pGalkB (Table 1) was sequenced and digested BamHI and NdeI to follow ligation in frame with tipAp of pIJ8600. The pIJalkB derivative (Table 1) were isolated and then transformed into *E.coli* ET12567 by electroporation. The ETpIJalkB was used to transform *S.coelicolor* M145 by coniugation according to *Kieser et al*., 2000. The resulting exconiugants were selected on apramycin. To evaluate the correct *alkB* gene expression in *S.coelicolor* M145, total RNA was extracted after thiostreptone induction in rich medium at 30°C, and used in RT-PCR assay as described below.

3.2. Analysis of the metabolic intermediates from *n*-alkanes oxidation pathway.

The metabolic intermediates resulting from incubation on C_{16} and C_{30} of M145- were analyzed by solid-phase micro extraction (SPME) coupled with GC-MS. *S.coelicolor* M145-AH and *S.coelicolor* M145 carrying the empty pIJ8600 were grown in JM medium, washed and re-suspended in BH medium to a final OD_{600} of 1.0 as previously described. The entire suspensions were analyzed by immersing the SPME fiber as previously described in chapter 1.

3.3. *In vivo* analysis of heterologous expression of *alkB*

The cells were broken by using 3 mg of lysozyme/ml in P-buffer (Kieser et al., 2000) and total RNA was isolated follow the same procedures described in chapter 1. Reverse transcription-PCR (RT-PCR) was performed by using the Superscript One-Step RT-PCR kit (Invitrogen) with about 0.1 µg of total RNA as a template, primer pairs internal to *alk* genes (AHqRT) (Table 3) and the conditions indicated by the supplier, routinely using 35 PCR

cycles. For each reaction, a negative control with *Taq* polymerase and without reverse transcriptase was included, in order to exclude DNA contamination.

3.4. Total protein extraction and DIGE analysis

For protein extraction, frozen biomass samples collected after different incubation time in mineral medium supplemented with hexadecane as sole carbon source, were sonicated according to Puglia et al. 1995. After dialysis against distilled water at 4°C and acetone precipitation at -20°C, proteins were suspended in IEF buffer containing 8M urea, 4% w/v CHAPS and 1% w/v 1,4-dithioerythritol (DTE) and stored at -80°C until use. For the first dimension, pH 4–7 18-cm IPG strips were rehydrated for 1 h in 340 µL of DeStreak rehydration solution (GE Healthcare, USA) containing 40 µg of sample proteins, 0.5% w/v IPG buffer (GE Healthcare) and 1% w/v DTE. Ettan IPGphor II system (GE Healthcare) was used for protein separation performing a 30 V pre-step for 10 h, followed by a 69 600 V-h IEF, with a maximum voltage of 8000 V. All the steps were performed at 20°C using 50 mA per strip. After IEF, the IPG strips were saturated with an equilibration buffer containing 6M urea, 30% v/v glycerol, 2%w/v SDS, 0.05M Tris-HCl, pH 6.8, and 2%w/v DTE for 12 min. Then, thiol groups were blocked by substituting DTE with 2.5% w/v iodoacetamide in the equilibrating buffer. The focused proteins were then separated on 12.5% polyacrylamide gels (SDS-PAGE) at 10°C in a Hoefer Dalt vertical system (GE Healthcare), using a maximum setting of 40 mA per gel and 110 V, according to manual (Amersham CyDye DIGE Fluors for Ettan DIGE)

Gels were matched using reproducible landmarks. The differential analysis was carried out on the biomasses collected from three parallel cultures. For each biomass, three technical gel replicates were performed and included for statistical analysis. Thus, protein spot quantification was calculated as mean %Vol (MV) by using data from six 2-DE gels. Mean-

squared deviation (MSD) was also calculated for statistical validation, according to software manufacturer's instructions.

4. Results

4.1. Experimental design

Genetic engineering strategies can be aimed to characterize metabolic pathways involved in long chain *n*-alkane utilization by gram positive bacteria. In fact, the development of high-throughput techniques to study global gene expression created new opportunities to identified proteins whose activity could be required to survive *n*-alkane-polluted environments. In particular, functional genomic and proteomic studies are a valuable tools to reveal gene expression profile changes during growth, generally portrayed as sets of genes being up- or down-regulated. These changes give clues to the regulation of many biological processes and metabolic pathways that allow the generation of metabolism intermediates. However, an approach based on global proteomic analysis cannot be applied to many alkane degrader strains since genome sequence have to be available. The unique genome sequence known for *Gordonia* genus is that of *G. bronchialis* (Ivanova et al., 2010). Hence, in order to characterize the metabolic pathways involved in long chain *n*-alkane utilization by gram positive bacteria, an highthroughput proteomic screening was carried out using the recombinant strain *S. coelicolor* M145-AH that express the *Gordonia* SoCg alkane monoxygenase (*alkB*). In fact, *S. coelicolor* is the best characterized actinomycete, used as model organism for secondary metabolite biosynthesis study. Its genome was sequenced so far (Bentley et al., 2002), several proteomics study were carried out to understand Streptomycete physiology and cellular differentiation and. In addition *S. coelicolor* M145-AH strain revealed to be a good microbial system for rapid *n*-hexadecane bioconversion to 1-hexadecanol and is also able to use this metabolic intermediate as carbon source (Lo Piccolo et al., 2010). Thus, this strain could be a good system to characterize *n*-alkane pathway degradation.

In order to identified proteins whose activity is required for growing on *n*-alkane, a proteomics analysis, based on 2D-PAGE and Mass spectrometry approaches, was performed on *S. coelicolor* M145-AH. This strain was incubated in three conditions using the mineral medium (BH) containing glucose (G) or hexadecane (H) as the sole carbon source or without any carbon source (0), respectively.

Proteomes of samples collected after 6h, 24h and 48h of incubation from each condition were analyzed by 2D-Differential Gel Electrophoresis (2D-DIGE). For isoelectrophocusing pH gradient ranging from 4 to 7 was used since previous study revealed that most *S. coelicolor* proteins possess an acid pI. In this study, protein spots were considered differentially abundant (up- or down-regulated) if their volume increased or decreased at least 1.5 fold, with a probability (p) for null hypothesis minor than 0.05 (ANOVA test).

In particular, protein spots showing no difference in abundance or being up- or downregulated in H in the respect of 0 cultivation were assigned to C_0 (for Constant), I_0 (for Increased) and D_0 (for Decreased) expression profiles, respectively, while protein spots showing no difference in abundance or being up- or downregulated in H in the respect of G condition were assigned to C_G, I_G or D_G expression profiles, respectively.

Since the objective of this study was to identify proteins specifically induced by the presence of *n*-alkane, the comparison between proteomes from H and 0 conditions during the first 48 h of growth was performed to reveal proteins whose expression is associated to *n*-alkane metabolic adaptation. On the other hand, the comparative analysis between H and G was carried out to asses if the expression of those proteins were specifically regulated by *n*-alkane metabolism or, in general, by carbon source utilization. Thus, a such similar proteome comparative analysis, with three levels of complexity, allowed to focus on the expression changes depending on physiological adaptations to the media.

4.2. General results

Surprisingly, up to 24 h H and 0 condition proteomes did not show any difference in protein pattern with the exception of 2 protein spots downregulated at 6 h and 24 h and 1 protein spots down- and upregulated at 24 h, respectively. On the contrary, 43 and 6 protein spots showed I_0 and D_0 expression profile, respectively, at 48 h. These result indicated that only after 24 h of incubation *n*-hexadecane presence causes changes in the expression profile. Thus, to reveal metabolic pathways (or part thereof) activated during n-alkane metabolism, only protein spots showing I_0 or D_0 expression profiles were identified by MS-MALDI-TOF analysis (Tab. 1).

On the other hand, the comparison of H and G proteomes revealed that, up to 6 h, 31 e 61 protein spots show I_G and D_G expression profile. At 24 h, the number of differentially abundant spots increased, with 32 and 133 proteins showed I_G and D_G expression profile, respectively. Due to the fact that H and 0 are essentially the same up to 24 h, these differences are not depending on *n*-hexadecane and glucose utilization but are likely to be depending exclusively on glucose metabolism. At 48 h 35 and 135 protein spots had I_G and D_G expression profile, respectively. Among them, a total number of 23 spots showed also either I_0 or D_0 profile at 48 h (Tab. 1), thus revealing a carbon source-depending expression profile.

4.3. Protein identification

The proteins differentially abundant at 48 h in H and 0 conditions were all identified and their expression profiles was monitored throughout time. This analysis revealed that at 48 h most of these proteins showed I_0 expression profile (88%), while only 12% had a D_0 profile (Fig3 A). Instead, after 6 h and 24 h of incubation, most of the identified spots (about 95%) had a C_0 profile and only few (about 5%) showed D_0 (Fig.3 A).

The expression profile of the identified proteins was monitored in G condition and the comparative analysis revealed that most of these proteins show a C_G profile. The number of C_G proteins decreased from 6 to 24 h and then remain nearly constant up to 48 h. Interestingly, the proteins showing a D_G profile increased from 26% to 40%, from 6 h to 24 h, respectively, while decreased (20%) at 48h (Fig 3B). The overall similarity between the H and 0 proteomes up to 24 h suggests that a similar metabolic *status* occurs in both conditions during the first growth phase. Thus, the differences between the H and the G proteomes up to 24 h are likely to be dependent exclusively on rapid utilization of glucose suggesting that *n*-hexadecane does not affect physiology up to 24 h and that cells need time to respond to 1-hexadecanol. In fact, only at 48 h proteome changes between H and 0 conditions reveal physiological adaptation to carbon source utilization.

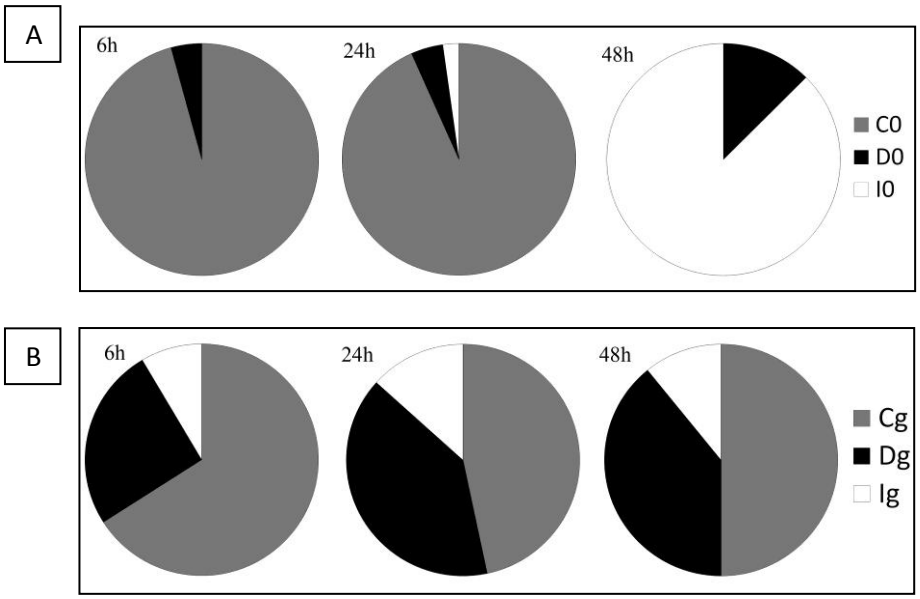

Figure 3. Temporal percentage distribution of expression profiles of proteins identified by MS analysis. A) comparative analysis between H and 0; B) comparative analysis between H and G

All 48 identified proteins were then classified based on their putative functions in 8 functional groups including glycolysis/gluconeogenesis, TCA, pentose phosphate pathway, fatty acid metabolism, amino acid and protein metabolism, respiration, transport and binding proteins (Table 1).

The expression profile of the identified proteins was combined between the two comparative analyses leading to the construction of a comprehensive set of 9 expression profiles (I_0Ig, I_0Dg, I_0Cg, D_0Ig, D_0Dg, D_0Cg, C_0Ig, C_0Dg and C_0Cg), that summarize the pattern of protein expression at three time points for each functional group. Then, common regulatory inferences were grouped in four pattern of profiles (Fig.4). Firstly, C_0Ig and C_0Dg profiles may be related to metabolic pathways or regulatory mechanism specifically concerning glucose utilization like glycolysis or catabolite repression (A pattern). The second domain concerns the proteins showing I_0Dg or D_0Ig are intermediate profiles that represent proteins whose expression depends on carbon source utilization, where glucose affects expression more than hexadecane does (B pattern). I_0Cg and D_0Cg profile which could be generally involved in responses due to any carbon source availability/utilization (C pattern). Finally, I_0Ig and D_0Dg profiles reflect an *n*-hexadecane-dependent expression (D pattern), thus, the presence of hexadecane specifically affects the expression of proteins with one of those profile.

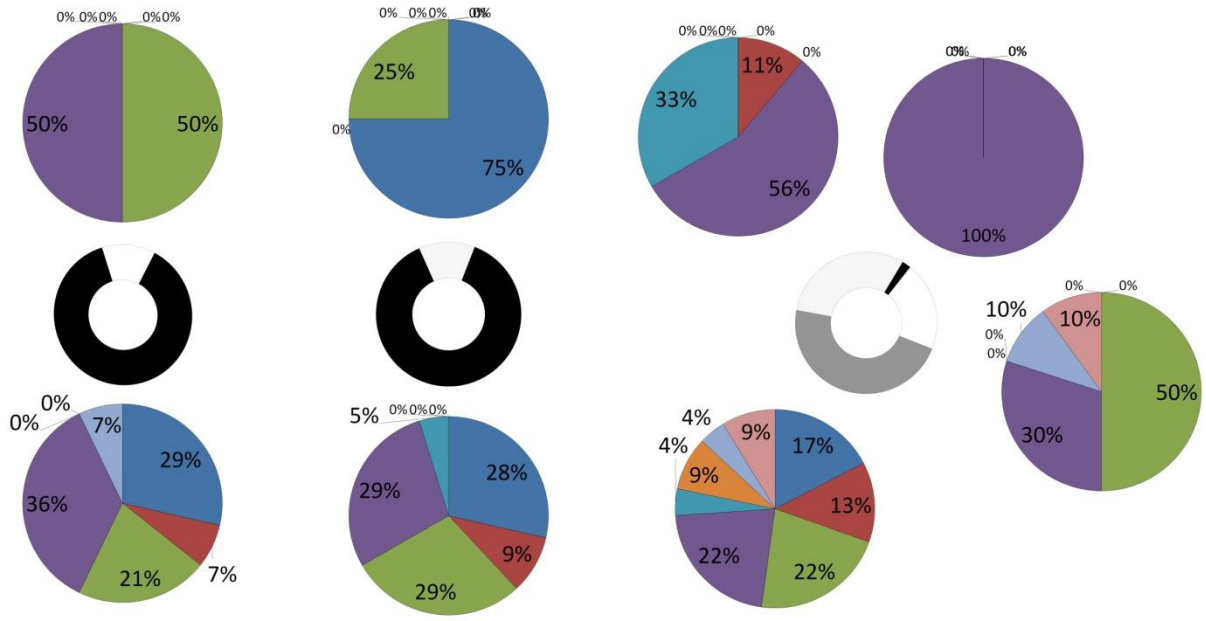

Figure 4. Biological inference of metabolic adaptation during growth. Distribution of common regulatory group at 6, 24 and 48 h (rings). Distribution of functional groups among the common regulatory group (pie chart).

The A pattern (black ring; metabolic pathways or regulatory mechanism specifically concerning glucose utilization) was the largest group of proteomics analysis at 6 and 24 h, mainly composed of proteins belonging to the functional group of the "Cental Carbon Metabolism", "Protein Metabolism" and "Transport and Binding Protein", thus reflecting glucose utilization. At 48h, in contrast, the number of spots belonging to A pattern significantly decreased and it was exclusively composed by proteins of the functional group of the "Protein Metabolism". In agreement with the activation of *n*-hexadecane metabolism, the B pattern (light gray ring; intermediate profiles that represent proteins whose expression depends on carbon source utilization where glucose affects expression more than *n*-hexadecane does) was present only at 48h. It was composed exclusively of proteins of the "Protein metabolism", "Amino acid metabolism" and "Respiration" groups. The C pattern (dark gray ring; generally involved in responses due to any carbon source

availability/utilization) appeared from 24 h and increased up to 48 h. If at 24 h it was essentially composed by protein of "Central Carbon Metabolism" group, it drastically becomes very heterogeneous at 48 h. These last results suggest that the activation of *n*-hexadecane metabolism at 48 h affects the expression of enzymes whose activity is related to generation of metabolic intermediates in catabolic routes and with the synthesis of cellular building blocks to support growth. In this context, the up- or downregulation of some proteins can be controlled by carbon source utilization where glucose is the stronger nutritional factor affecting growth. The D pattern (white ring; profiles reflect an *n*-hexadecane-dependent expression) was poorly represented at 6 h and completely disappeared at 24 h, but it becomes one of the main groups at 48 h. It was essentially composed by proteins of the "Transport and Binding Proteins" and "Proteins Metabolism" groups over time. These proteins are mostly membrane proteins or involved in protein synthesis, suggesting that the adaptation to *n*-alkane metabolism could be driven by membrane alteration and alter protein modification/synthesis.

5. Main remarks

The main concern of this study revealed that at least 24 h are necessary for *n*-hexadecane utilization in *S. coelicolor* M145 AH, although 1-hexadecanol formation is revealed after 4 h. The comparative proteome analysis between H and 0 conditions revealed that the differentially expressed proteins, identified by MS analysis, are mainly upregulated (Fig. 3) and involved in Central Carbon and Protein Metabolism (Fig. 5), thus reflecting the activation of metabolic routes to support growth.

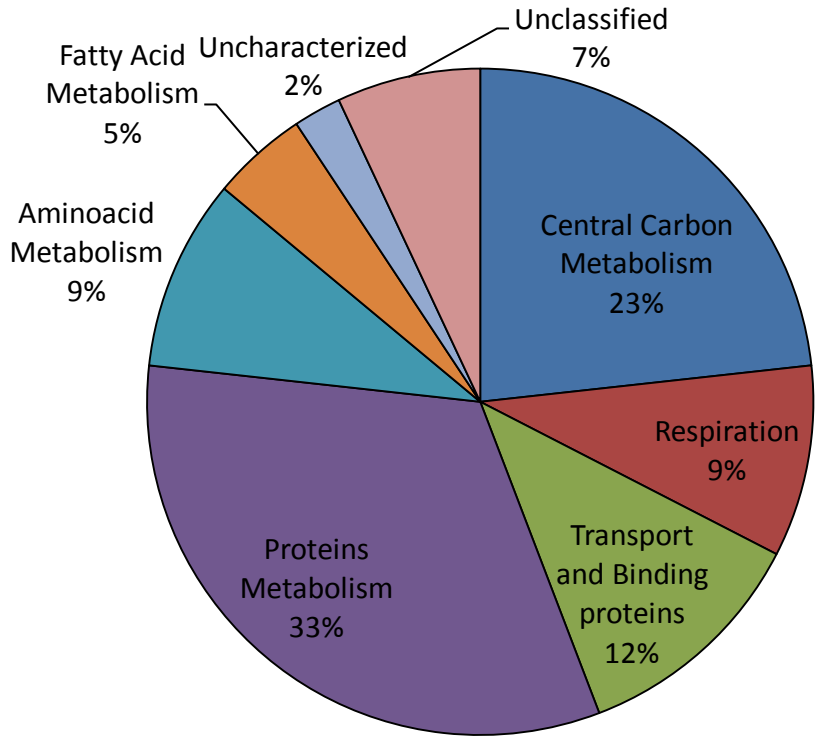

Figure 5. Schematic representation of composition of proteins that showed I_0 expression profile at 48 h.

In agreement with this last evidence, the downregulated proteins are reported to be associated with nutritional stress limitation/starvation (ref), thus reflecting the absence of carbon source in 0. The comparative expression profile of these proteins between H and G conditions revealed that no one was a unique protein, but different levels of protein abundance could differentially regulate cellular activities and metabolic intermediate distribution indicating that *n*-alkane was metabolized via pathways similar to those involved in glucose metabolism.

At 48 h, among the I_0 proteins, those involved in the "Central Carbon Metabolism", "Transport and Binding Proteins" and "Aminoacids and Proteins Metabolism" showed D_G profile, revealing a strong control by glucose. Thus, these results indicate that even if *n*-

hexadecane metabolism was activated, the expression of these proteins is higher during glucose utilization. An interesting exception are proteins involved in fatty acid metabolism that have I_0C_G profile at 48 h. It was demonstrated by Banchio C. and Gramajo H.C., (see introduction) that enzymes of the β-oxidation cycle are not inducted by substrates. Consistent with this, this analysis showed that proteins of the "Fatty acid metabolism" group had C_0C_G expression profile at 6 and 24 h (Table 1), probably because their expression drastically reduced at 48 h, when the carbon source absence was sensed in 0. Thus, their decreased expression in 0 could be a molecular signaling of the cell famine. Anyway, analysis of the cellular vitality coupled with this proteomic approach suggested that *S. coelicolor* significantly reduces its metabolism, but does not switch off it and surprisingly, remains viable.

Synoptic picture is represented by a metabolic network showing the connection between anabolic and catabolic pathways during time course of growth in the three conditions, highlighting enzymes that could play a key role in controlling metabolic intermediate flux (fig.6).

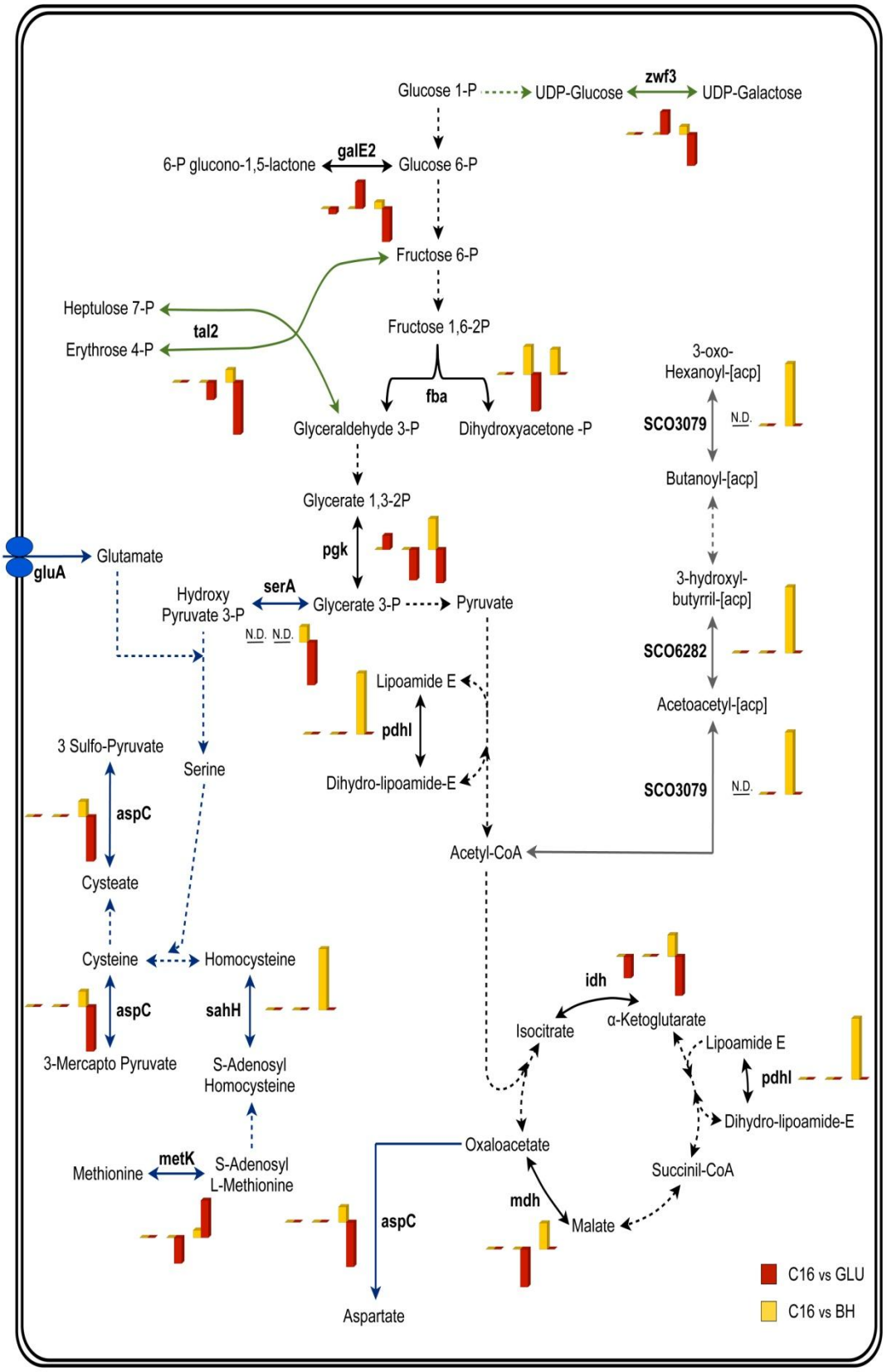

Figure 6. Synoptic picture of the metabolic network showing the connection between anabolic and catabolic pathways during time course of growth in the three conditions

Conclusions

Conclusions

Pollution of soil and water environments by crude oil has been, and is still today, an important problem. Crude oil is a complex mixture of thousands of compounds. Among them, alkanes constitute the major fraction. Alkanes are saturated hydrocarbons of different sizes and structures. Although they are chemically very inert, most of them can be efficiently degraded by several microorganisms.

Research performed in the last few years has resulted in many new insights on how microorganisms degrade alkanes. However, several aspects remain poorly understood. One is how alkanes are incorporated or transported into the cell, which may differ for different alkanes and for different microorganisms. The enzymes for the degradation of low- and medium-chain length alkanes are rather well characterized, except for the paucity of structural data. However, new and different enzymes have recently been found that oxidize high-molecular-weight (C20–C50) alkanes. There are several hints indicating that additional new alkane hydroxylases do exist that have still not been identified. It is also intriguing why bacterial strains frequently contain several alkane hydroxylases that have very similar substrate specificities. Perhaps these hydroxylases differ in aspects that are still unknown but that are important for cell biology. Regulation of the genes coding for the alkane degradation pathways still has many unresolved issues. It has been important to realize that the expression of the alkane-degradation pathways is frequently coordinated with other aspects of cell metabolism. Most efforts have been directed to elucidate the mechanisms responsible for the catabolite repression control, but there are other global regulation phenomena that modulate the expression of alkane-degradation genes. Elucidation of these mechanisms is important to understand how alkanes are degraded in Nature and to design bioremediation strategies that are efficient at stimulating the degradation of alkanes in contaminated sites.

The picture is far from clear and requires much more efforts in different microorganisms, because the molecular mechanisms will likely be different in each case.

Here I describe for the first time the unique functional AlkB-type alkane hydroxylase system that allows growth on long chain liquid and solid *n*-alkanes in the Gram-positive *Gordonia* strain SoCg. Up to date the only long chain alkane hydroxylase system of this genus that has been characterized is that of *Gordonia* TF6 that was found to be functional on *n*-alkanes from C_5 to C_{13} (Fujii et al., 2004).

The alkane hydroxylase from *Gordonia* SoCg is not only active on a wide range of long chain liquid and solid *n*-alkanes, but revealed to be also a more versatile enzyme than *Pseudomonas*-type *n*-alkane 1-monooxygenases as it is able to use other electron transfer systems in the absence of its two specific components, rubredoxin and rubredoxin reductase. This versatility seems to be shared by other phylogenetically close AH from strains belonging to the *Corynebacterium-Mycobacterium-Nocardia* (CMN) complex (Fujii et al., 2004; Sameshima et al., 2008) strengthening the interest towards this group of bacteria.

Although hydrocarbon utilization is quite common among bacterial strains belonging to the CMN complex, only a few strains, mainly rhodococci, have been studied in some details. SoCg is the first *Gordonia* strain that is able to grow on solid *n*-alkanes to be characterized.

It is known that when the carbon source is in excess relative to nitrogen, many bacteria transform part of this into storage materials such as triacylglycerols, wax esters, poly(hydroxybutyrate) or poly(3-hydroxyalkanoates), which accumulate as lipid bodies or as granules.

The recalcitrant liquid and solid *n*-alkane are oxidize to long chain fatty alcohol and finally, completely metabolize via oxidative process to fatty acid needed to support growth of the

cell. It was demonstrated (Ishige et al., 2003) that part of the 1-alkanol is esterified with fatty acid CoA thioesters by the enzyme WS/DGAT (Wältermann et al., 2007). This reaction generate high energy value molecules, that are packed in internal vesicles as storage compounds. Interestingly, *Gordonia sp.* strain SoCg is able to produce neutral lipids during growth in the presence of *n*-alkanes as the sole carbone source and in particolar, TAGs and Wes were identified as major products of growth on C_{16} and C_{18}.

This study evaluates the capacity for storage compound metabolism by *Gordonia sp*. strain SoCg and demonstrates the functionality of the biosynthetic pathways. This knowledge constitutes a framework for additional studies to determine the physiological and ecological functions and significance of these storage compounds for SoCg. The occurrence of diverse storage compounds in SoCg emphasizes the complexity of the physiology and biochemistry of this heterotrophic soil bacterium. This complexity reflects the richness, diversity and versatility of lipid metabolism in this strain and related lipid-rich bacteria. The ability of these bacteria to accumulate diverse storage compounds may permit cells to rapidly respond to stress and to maintain active its metabolism under fluctuating environmental conditions, providing them with an adaptive advantage over less versatile bacteria.

Beyond that, TAGs are excellent reserve materials for several reasons. Their extreme hydrophobicity allows their accumulation in large amounts in cells without changing the osmolarity of the cytoplasm. Oxidation of TAGs produces the relatively high yields of energy in comparison with other storage compounds such as PHA and carbohydrates, since the carbon atoms of TAG acyl residues are in a very reduced state (Alvarez et al., 2002). For this reason, neutral lipids are widely used in the manufacture of fine chemicals and pharmaceutical applications, as lubricants and food additives. Also, they are of interest in medical and cosmetic industries. Currently, most of Wes are isolated from plants, while the rest is produced via chemical synthesis or through immobilized lipases in bioreactors.

However, both of these procedures are very expensive, so further investigation on neutral lipids production by *Gordonia sp*. SoCg could provide cost effectiveness procedure to their use as biocatalysts.

Additional progress has been made in understanding the metabolism of *n*-alkanes in gram-positive bacteria. The proteomic analysis has identified a set of proteins specific induced by the presence of *n*-hexadecane in the recombinant *S.coelicolor* M145, expressing the *Gordonia* SoCg AlkB. Interestingly, it was revealed that the most of phisiological changes induced by the adaptantion to *n*-alkane metabolism are related to the membrane alteration and protein modification/synthesis. The main concern of this study revealed that at least 24 h are necessary for *n*-hexadecane utilization in *S. coelicolor* M145 AH, although 1-hexadecanol formation is revealed after 4 h. Surpraisingly, no unique proteins were detected, but different levels of protein abundance could differentially regulate cellular activities and metabolic intermediate distribution indicating that *n*-alkane was metabolized via pathways similar to those involved in glucose metabolism.

This results provide a new omics platform to understand global gene expression regulation required to survive in *n*-alkane-polluted environments.

Tables

TABLE 1. Bacterial strains and plasmids used in this study

Strains or vectors	Description	Reference or source
Strains		
Gordonia sp. SoCg	Long chain *n*-alkane degrader, *alkB*+	Quatrini et al., 2008
Gordonia Ω*alkB*	*Gordonia* sp. strain SoCg distruption mutant, *alkB*$^-$, apra^r	This study
Escherichia coli DH10B	F$^-$ mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL nupG.	Invitrogen
Escherichia coli ET12567	F$^-$, dam$^-$13::Tn9 dcm- 6hsdM hsdR, zjj-202::Tn10, recF143, galK2, galT22, ara14, lacY1, xyl-, 5leuB6 Cml^r, Tet^r, Kan^r.	Kieser et al., 2000
Escherichia coli BL21 (DE3) pLysS	F$^-$, *omp*T, *hsd*S$_B$ (r_B^-, m_B^-), *dcm*, *gal*, λ(DE3), pLysS, Cam^r.	Invitrogen
Escherichia coli BL21-AH	*E. coli* BL21 containing the recombinant expression vector pRalkB; *alkB*^{r} , amp^r	This study
Streptomyces coelicolor M145	wild type; SCP1$^-$, SCP2$^-$.	Kieser et al., 2000
Streptomyces coelicolor M145-AH	*S.coelicolor* M145 containing the recombinant integrative pIalkB; *alkB*^{r} , thio^r, apra^r	This study
Cloning and expression vectors		
pUC18	*E. coli* cloning vector, Amp^r	Invitrogen
pGEM®-T Easy vector	*E. coli* cloning vector, Amp^r	Promega
TOPO-TA®	*E. coli* cloning vector, Amp^r	Invitrogen
pRSET-B	*E. coli* expression vector, Amp^r	Invitrogen
pIJ8600	Streptomycete expression vector Apra^r Promotor induction by Thiostreptone	Kieser et al., 2000
pIJ773	Streptomycete cloning vector, Apra^r;used to extract apramycin resistance cassette *aac*(3)IV with its oriT	Kieser et al., 2000
Plasmids containing DNA from *Gordonia* SoCg		
palkCg23	pGEM®-T Easy vector derivative containing a 570 bp *alkB* fragment [Genbank acc.no.EF437969]	Quatrini et al., 2008
palk68	pUC 18 derivative, containing a 8 kb fragment of *Gordonia* SoCg including the *alk* cluster	This study
pGalkB1	pGEM®-T Easy vector derivative, containing *alkB* (1,2 kb), amplified by PCR with primers alkNIFor and alkBHIRev	This study
pRalkB	pRSET-B derivative, containing *alkB* (1,2 kb)	This study

pIalkB	pIJ8600 derivative, containing *alkB* (1,2 kb)	This study
palkapra	palk68 derivative, containing the apramycin resistance cassette cloned into an *Ale*I unique restriction site of *alkB*	This study
ptgs-int	TOPO-TA derivative, containing the internal region of putative *atfa* gene	This study

TABLE 2. Primers used in this study

Primer	Sequence	Reference or source
AH+FR	**5'**- Y RCS GCV CAC GAR YTS GGB CAC AAG-**3'**	Quatrini et al., 2008
AH+RV	**5'**- SGG ATT CGC RTG RTG RTC RCT GTG -**3'**	Quatrini et al., 2008
AHqRTFor	**5'**-GGACCGATGCTGGTCTATGT-**3'**	This study
AHqRTRev	**5'**-CAGATAACAGGCCATGACGA-**3'**	This study
rubA3qRTFor	**5'**-CTACCGTGTCCGGTCTGTG-**3'**	This study
rubA3qRTRev	**5'**-CCAGTCGTCGGGAATGTC-**3'**	This study
rubA4qRTFor	**5'**-CTGCGAGGTCTGCGGATT-**3'**	This study
rubA4qRTRev	**5'**-GGCCACCTCGACCATCTC-**3'**	This study
rubBqRTFor	**5'**-GGGTGTTGATCCAGTTCAGG-**3'**	This study
rubBqRTRev	**5'**-TATCTGGCACATCACCAACG-**3'**	This study
alkUqRTFor	**5'**-GCGTTCACCGAGTACTTCAC-**3'**	This study
alkUqRTRev	**5'**-ATCGACAACCACGTCGACTC-**3'**	This study
alkNIFor	**5'**-AACATATGCTCGTGAGAGGAGCGTGC-**3'**	This study
alkBHIRev	**5'**-AAGGATCCCCGGACAACGGTAGGCGC-**3'**	This study
CF	**5'**-ATGTTYATHGCNATGGAYCCNC-**3'**	Kubota et al., 2005
CR	**5'**-NARNCKRTTNCCCATRCANCKRTG-**3'**	Kubota et al., 2005
apra750FR	**5'**-ATTCCGGGGATCCGTCGACC-**3'**	This study
apra750RV	**5'**-TGTAGGCTGGAGCTGCTTC-**3'**	This study
ladAFR	**5'**-GGC GTS TAC GMC RWC TAC GGY RGG-**3'**	This study
ladARV	**5'**-GAY CTA CCA GGY CGG GTC GTC G-**3'**	This study
alkCG341FR	**5'**-CCG AGG ACC CGG CGA GCT C-**3'**	This study
alkCG341RV	**5'**-CTC CGG GGT GCA CCG CTC-**3'**	This study
tgs-int 1	**5'**-TCSCGCCCGCTSTGGGAG-**3'**	Alvarez et al., 2008
tgs-int 2	**5'**-SGGGCCSGGSACGTTCGA-**3'**	Alvarez et al., 2008

atfa FR	**5'**- GCACGTTCGAACGCCTCGGC -**3'**	This study
atfa RV	**5'**- CACTGAACTGCGGCACGACG -**3'**	This study
atfa2FR	**5'**- CTTCGACCTGCACTACCACA -**3'**	This study
atfa2RV	**5'**- GCCCAGCTTCTCGTGATACC -**3'**	This study
diacylFR	**5'**- TCACCCGGCTTCCCTGGCCAG -**3'**	This study
diacylRV	**5'**- ATGACCGGACCCGACGCGCTG -**3'**	This study
4163FR	**5'**- ATGCCATACATGCCGGTCACG -**3'**	This study
4163RV	**5'**- CGGGCCGACGTGGGCCAC -**3'**	This study
3771FR	**5'**- TCAATCGCAGGCGGCGAGTA -**3'**	This study
3171RV	**5'**- ATGGAACGACTGAGCGGACTG -**3'**	This study

TABLE 3. Genes identified and sequence similarities in *Gordonia* sp. strain *SoCg alk* locus.

Gene	Length (aa)	Best BLASTP Aligment	Overlap/Identities
orf1	124	conserved hypothetical protein [*Rhodococcus equi* ATCC 33707] [acc.no.06829834.1]	48/124 (38%)
alkB	411	alkane-1-monooxygenase [*Gordonia* sp. TF6] [acc.no. BAD67020.1]	287/323 (88%)
rubA3	55	rubredoxin 3 [*Gordonia sp. TF6*] [acc.no. BAD67021.1]	45/54 (83%)
rubA4	61	rubredoxin 4 [*Gordonia sp. TF6*] [acc.no. BAD67022.1]	34/59 (57%)
rubB	400	rubredoxin reductase [*Gordonia* sp. TF6] [acc.no BAD67023.1]	249/349 (71%)
alkU	160	putative transcriptional regulator, TetR family [*Mycobacterium abscessus*] [acc.no YP 001704325.1]	87/168 (51%)